

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2003, 01:06:07 ; Search time 94 Seconds  
(without alignments)  
5367.031 Million cell updates/sec

Title: US-09-508-997A-1  
Perfect score: 1143  
Sequence: 1 atggctcccttgctgcgcg.....gcaagattcttgccottga 1143

Scoring table: OLIGO NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCUTS\_COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1143	100.0	1950	2	US-08-472-659-30
2	1143	100.0	1950	2	US-08-474-661-30
3	1143	100.0	1950	2	US-08-611-977-30
4	254	22.2	254	4	US-09-016-434-658
5	230	20.1	316	4	US-09-016-434-824
6	121	10.6	600	4	US-09-702-705-233
7	121	10.6	600	4	US-09-736-457-233
8	25	2.2	37	2	US-08-472-659-32
9	25	2.2	37	2	US-08-474-661-32
10	25	2.2	37	2	US-08-611-977-32
11	23	2.0	23	2	US-08-472-659-14
12	23	2.0	23	2	US-08-474-661-14
13	23	2.0	23	2	US-08-611-977-14
14	21	1.8	21	2	US-08-472-659-15
15	21	1.8	21	2	US-08-472-659-16
16	21	1.8	21	2	US-08-472-659-27
17	21	1.8	21	2	US-08-472-659-28
18	21	1.8	21	2	US-08-472-659-29
19	21	1.8	21	2	US-08-474-661-15
20	21	1.8	21	2	US-08-474-661-16
21	21	1.8	21	2	US-08-474-661-27
22	21	1.8	21	2	US-08-474-661-28
23	21	1.8	21	2	US-08-474-661-29
24	21	1.8	21	2	US-08-611-977-15
25	21	1.8	21	2	US-08-611-977-16
26	21	1.8	21	2	US-08-611-977-27
27	21	1.8	21	2	US-08-611-977-28

Sequence 29, Appl  
Sequence 3, Appl  
Sequence 150, App  
Sequence 291, App  
Sequence 221, Ap  
Sequence 57, Appl  
Sequence 24, Appl  
Sequence 10, Appl  
Sequence 1, Appl  
Sequence 3, Appl  
Sequence 11, Appl  
Sequence 460, App  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 252, App  
Sequence 402, App

ALIGNMENTS

RESULT 1  
US-08-472-659-30  
; Sequence 30, Application US/08472659  
; Patent No. 5831030  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Fuyuki  
; APPLICANT: TSUJIMOTO, No. 5831030uo  
; APPLICANT: NAKAZATO, Hiroshi  
; APPLICANT: MURA, Kenju  
; APPLICANT: ISHIDA, No. 5831030uhiro  
; APPLICANT: KURIHARA, Tatsuya  
; APPLICANT: YAMAUCHI, Kozi  
; APPLICANT: YAMAGUCHI, No. 5831030oml  
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,659  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 4-212305  
; FILING DATE: 17-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-067339  
; FILING DATE: 04-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/091,028  
; FILING DATE: 14-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McGowan, Malcolm K.  
; REGISTRATION NUMBER: 39,300  
; REFERENCE/DOCKET NUMBER: 001560-248  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:

Query Match	100.08;	Score 1143;	DB 2;	Length 1950;
Best Local Similarity	100.04;	Pred. No. 0;		
Matches 1143;	Conservative	0;	Mismatches	0;
			Indels	Gaps
QY	1	ATGGCTCCCTCTGCTGCAGCAAAATGACAGAGTATTTTCCTTCAAACTGTTACAGAGATGAT	60	
Db	74	ATGGCTCCCTTGCCTGCAGCAAAATGACAGAGTATTTTCCTTCAAACTGTTACAGAGATGAT	133	
QY	61	GACAAATCAGGAANAATGGAAATATGTTCTTTCTCTGAGGCTCTTGCTGCTGCTGCTGCTG	120	
Db	134	GACAAATCAGGAANAATGGAAATATGTTCTTTCTCTGAGGCTCTTGCTGCTGCTGCTGCTG	193	
QY	121	CTGGTCCGCTTGGGCGCTCMAGATGACTCCCTCTCTCAGATGATAAGTTGCTTCAATGTT	180	
Db	194	CTGGTCCGCTTGGGCGCTCMAGATGACTCCCTCTCTCAGATGATAAGTTGCTTCAATGTT	253	
QY	181	AACACTGGCTCAGGATATCGAAACTCTTTTAAATATGATGTCGGGCTCCAGTCTCAACTG	240	
Db	254	AACACTGGCTCAGGATATCGAAACTCTTTTAAATGATGTCGGGCTCCAGTCTCAACTG	313	
QY	241	AAAAGAGTTTTTCTGATATTAATGATATCCCAAGAGATATGATCTCAGCATTTGTAAT	300	
Db	314	AAAAGAGTTTTTCTGATATTAATGATATCCCAAGAGATATGATCTCAGCATTTGTAAT	373	
QY	301	GGGCTTTTGTGTAAGAAAGTATGCGCTTTTCATPAAGGACTACATTGAGTGTGCCGAAAA	360	
Db	374	GGGCTTTTGTGTAAGAAAGTATGCGCTTTTCATPAAGGACTACATTGAGTGTGCCGAAAA	433	
QY	361	TTATPACATGCCAAAGTGGAGCGAGTTGCACTTTTACGATCATTTAGAGACACTAGACGT	420	
Db	434	TTATPACATGCCAAAGTGGAGCGAGTTGCACTTTTACGATCATTTAGAGACACTAGACGT	493	
QY	421	AAATATTAATAGCTGGTTGAAATGAACAACATGCCAAATCAAGAACCTGATTTGGTGAA	480	
Db	494	AAATATTAATAGCTGGTTGAAATGAACAACATGCCAAATCAAGAACCTGATTTGGTGAA	553	
QY	481	GGTGGCAATAGCTCATCTGCTGTATAGTGTGCTGGTGAATGCTGTGTACTTCAAAGCGAG	540	
Db	554	GGTGGCAATAGCTCATCTGCTGTATAGTGTGCTGGTGAATGCTGTGTACTTCAAAGCGAG	613	
QY	541	TGGCAATCAGCTTCCACAGAGCGAAACCAATTAATTTGCCATTTCAAACTCTCCCAAGTGC	600	
Db	614	TGGCAATCAGCTTCCACAGAGCGAAACCAATTAATTTGCCATTTCAAACTCTCCCAAGTGC	673	
QY	601	TCGTGGAGGACGCTGGCCATGATGCATCAGGAAACGGAAGTTCAATTTGCTGTATTAG	660	
Db	674	TCGTGGAGGACGCTGGCCATGATGCATCAGGAAACGGAAGTTCAATTTGCTGTATTAG	733	
QY	661	GACCCATCAATGAAGATCTTTGAGCTCAGATACAAATGGTGGCATTAACATGTACCTCTG	720	
Db	734	GACCCATCAATGAAGATCTTTGAGCTCAGATACAAATGGTGGCATTAACATGTACCTCTG	793	
QY	721	CTGGCTGAGATGACCTCTCTGAAATGGAATGAACAACTGACCTTTCAGATCTTAATGGAA	780	
Db	794	CTGGCTGAGATGACCTCTCTGAAATGGAATGAACAACTGACCTTTCAGATCTTAATGGAA	853	

RESULT 2  
US-08-474-661-30  
Sequence 30, Application US/08474661  
Patent No. 5874253  
GENERAL INFORMATION:  
APPLICANT: TSUTSUMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSURUOKA, No. 5874253uo  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. 5874253uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAICHI, Koza  
APPLICANT: YAMAGUCHI, No. 5874253omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESS: Burns, Doane, Swecker & Mathis  
STREET: George Mason Bldg., Washington & Prince Sts.  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,661  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-067339  
FILING DATE: 04-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: REA, TERESA STANEK  
REGISTRATION NUMBER: 30,427

REFERENCE/DOCKET NUMBER: 001560-204

REFERENCE/DOCID NUMBER: 00  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-6620

TELETYPE. (703) 536 0020  
INFORMATION FOR SEQ ID NO: 30:

### SEQUENCE CHARACTERISTICS:

LENGTH: 1950 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

STRAIN: A431

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1

OTHER INFORMATION: /note= "DNA coding for human

OTHER INFORMATION: megakaryocyte differentiation factor."

FEATURE:

NAME/KEY: CDS

LOCATION: 74..1217

US-08-474-661-30

Query Match 100.0%; Score 1143; DB 2; Length 1950;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCTCCCTTCTGCTCGACCAATGACAGAGTTTGCCTTCAACCTGTTACAGAGAGATGGAT	60
DB	74	ATGGCTCCCTTCTGCTCGACCAATGACAGAGTTTGCCTTCAACCTGTTACAGAGAGATGGAT	133
QY	61	GCAATCAAGGAAATGGAATGTGTTCTTTTCTCTCTCTGAGCCTCTTCGCTGGCCCTGGCC	120
DB	134	GACATCAAGGAAATGGAATGTGTTCTTTCTCTCTGAGCCTCTTCGCTGGCCCTGGCC	193
QY	121	CTGGTCCGCTTGGCGCTCAGATGACTCCCTCTCTCAGATTGATGAAGTTCGTTCAATGTT	180
DB	194	CTGGTCCGCTTGGCGCTCAGATGACTCCCTCTCTCAGATTGATGAAGTTCGTTCAATGTT	253
QY	181	AACACTGCTCAGGATATGGAACCTTTCTTAATAGTCAGTCAGGCGTCCAGTCTCAACTG	240
DB	254	ACACTGCTCAGGATATGGAACCTTTCTTAATAGTCAGTCAGGCGTCCAGTCTCAACTG	313
QY	241	AAAGAGTTTTTCTCGATAAAATGCATCCCAAGGATTATGACTCAGCATTTGTGAAT	300
DB	314	AAAGAGTTTTTCTCGATAAAATGCATCCCAAGGATTATGACTCAGCATTTGTGAAT	373
QY	301	GGGCTTTTCTGAAAAGAGTGTGGCTTTCTAATAGACTACATTTAGTGTGGCGAAAAA	360
DB	374	GGGCTTTTCTGAAAAGAGTGTGGCTTTCTAATAGACTACATTTAGTGTGGCGAAAAA	433
QY	361	TTATAGATGCAAGTGGAGCGAGTTGACTTTACGAATCATTTAGAAGACACTTAGACGT	420
DB	434	TTATAGATGCAAGTGGAGCGAGTTGACTTTACGAATCATTTAGAAGACACTTAGACGT	493
QY	421	AATATTATTAAGTGGTTCGAATAAACAACATGCGCAAAATCAAGAGAGTCATTTGTGAA	480
DB	494	AATATTATTAAGTGGTTCGAATAAACAACATGCGCAAAATCAAGAGAGTCATTTGTGAA	553
QY	481	GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGGTGAATGCTGTACTTCAAAAGGCAAG	540
DB	554	GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGGTGAATGCTGTACTTCAAAAGGCAAG	613
QY	541	TGGCAATCAGCGCTCACCAGAGCGAAACCATAAATTCGCCATTTCAAATCTCCCAAGTGC	600
DB	614	TGGCAATCAGCGCTCACCAGAGCGAAACCATAAATTCGCCATTTCAAATCTCCCAAGTGC	673
QY	601	TCGTGGAGAGGAGTCGCCATGATCATCAGGAAACGGAAGTTCAATTTGCTGTTATTAG	660
DB	674	TCGTGGAGAGGAGTCGCCATGATCATCAGGAAACGGAAGTTCAATTTGCTGTTATTAG	733
QY	661	GACCCATCAATGAAGATCTCTGAGCTCAGATACAATGTGGGCAATAACAATGATCGTTCTG	720

Query Match	100.0%	Score 1143	DB 2	Length 1950
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1143	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGCCCTCCCTTGCCTCAGCAAAATG	CAGAGTTTTGCTTCAACCTGTTCA	GAGAGATGGAT 60
DB	74	ATGGCTTCCTTGCCTCAGCAAAATG	CAGAGTTTTGCTTCAACCTGTTCA	GAGAGATGGAT 133
QY	61	GACAAATCAAGGAAATGGAATGTGTTCTTTCCTCTCTGAGCTCTTGGTGCCTGGCC	120	
DB	134	GACAATCAAGGAAATGGAATGTGTTCTTTCCTCTCTGAGCTCTTGGTGCCTGGCC	195	
QY	121	CTGGTCGCTTGGCGCTCAGAGTACTCCCTCTCTCAGATTGATAAGTGTCTCAATGTT	180	
DB	194	CTGGTCGCTTGGCGCTCAGAGTACTCCCTCTCTCAGATTGATAAGTGTCTCAATGTT	253	
QY	181	AACACTGCCTCAGGATATGGAATACTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG	240	
DB	254	AACACTGCCTCAGGATATGGAATACTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG	313	
QY	241	AAAGAGTTTTTCTGATATAAATGCAATCCCAACAGATTATGATCTCAGCATTTGTAAT	300	
DB	314	AAAGAGTTTTTCTGATATAAATGCAATCCCAACAGATTATGATCTCAGCATTTGTAAT	373	
QY	301	GGCGTTTTTCTGAAATAAGTGTATGGCTTTCATAGGACTACATTGAGTGTGCGGAAAAA	360	
DB	374	GGCGTTTTTCTGAAATAAGTGTATGGCTTTCATAGGACTACATTGAGTGTGCGGAAAAA	433	
QY	361	TTATACGATGCCAAAGTCGAGCGAGTTGACTTTACGAATCAATTAGACAGACTTAGACGT	420	
DB	434	TTATACGATGCCAAAGTCGAGCGAGTTGACTTTACGAATCAATTAGACAGACTTAGACGT	493	
QY	421	AATATTAATAAGTGGGTTCGAAATGAACACATGCAAAATCAAGAACTGATTGGTGA	480	
DB	494	AATATTAATAAGTGGGTTCGAAATGAACACATGCAAAATCAAGAACTGATTGGTGA	553	
QY	481	GGTGGCATAAGCTCATCTGCTTAATGCTGCTGGTGAATGCTGTGTACTTCAAAGGCAAG	540	
DB	554	GGTGGCATAAGCTCATCTGCTTAATGCTGCTGGTGAATGCTGTGTACTTCAAAGGCAAG	613	
QY	541	TGGCAATCAGCCTTCCAAAGAGCGAAACCAATTAATTCGCAATTTCAANTCTCCCACTGC	600	

RESULT 4  
US-03-016-434-658  
; Sequence 658, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016.434  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:



Query Match 20.1%; Score 230; DB 4; Length 316;  
Best Local Similarity 99.6%; Pred. No. 6.3e-110;  
Matches 280; Conservative 0; Mismatches 1; Indels

Qy	822	ATTTTTCCTCAGTTCAGATTCAGAGAGAAATTAAGAAATCGAACAAATATTTGAGAGCCCT	881
Db	1	AGTTTTTCCCTCAGTTCAGATTCAGAGAGAAATTAAGAAATCGAACAAATATTTGAGAGCCCT	50
Qy	882	AGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCGGGATGGCTTCGGGGGG	941
Db	61	AGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGTTGGTTCGGGGGG	120
Qy	942	TGCTCTGTATATATCAAGGATGATGCACAAATCTTACATAGAGGTCACTGAGAGGGGCAAC	1001
Db	121	TGCTCTGTATATATCAAGGATGATGCACAAATCTTACATAGAGGTCACTGAGAGGGGCAAC	180
Qy	1002	CGAGCTACTGCTGGCACAGGAAGTAATATTCTGAGAAAAGCAACTCCCTCAGTCCACGCT	1061
Db	181	CGAGCTACTGCTGGCACAGGAAGTAATATTCTGAGAAAAGCAACTCCCTCAGTCCACGCT	240
Qy	1062	GTTTAGAGCTGACACCCCATTTCTTATTGTTATCAGGAGG	1102
Db	241	GTTTAGAGCTGACACCCCATTTCTTATTGTTATCAGGAGG	281

```

RESULT 6
US-09-702-705-233/c
; Sequence 233, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-233
Query Watch: 10.6%; Score 121; DB 4; Length 600;
Best Match Similarity 100.0%; Pred.No.3.7e-53;
Local 121; Conservative 0; Mismatches 0; Indels 0;

```

```

Query Match: 10.6%; Score 121; DB 4; Length 600;
Best Local Similarity 100.0%; Pred.No. 3.7e-53;
Matches 121; Conservative 0; Mismatches 0; Indels

```

Mon Dec 15 08:51:17 2003

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,659

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 001560-248

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 37 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-472-659-32

Query Match 2.2%; Score 25; DB 2; Length 37;

Best Local Similarity 100.0%; Pred. No. 0.0032;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTCCCTTCCTGCGCAATG 25

Db 13 ATGGCTCCCTTCCTGCGCAATG 37

RESULT 9

US-08-474-661-32

Sequence 32, Application US/08474661

Patent No. 5874253

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5874253uo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5874253uhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAUCHI, Kozo

APPLICANT: YAMAGUCHI, No. 5874253omi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: George Mason Bldg., Washington & Prince Sts.

CITY: Alexandria

STATE: Virginia

Qy 1023 AAGTATATTGTAGAAAGCACTCCCTCAGTCACGCTGTTTAGAGCTGACCAACCAATT 1082

Db 600 AAGTATATTGTAGAAAGCACTCCCTCAGTCACGCTGTTTAGAGCTGACCAACCAATT 541

Qy 1083 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142

Db 540 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481

Qy 1143 A 1143

Db 480 A 480

RESULT 7

US-09-736-457-233/c

Sequence 233, Application US/09736457

Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary

APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darrick

APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane

APPLICANT: Fan, Liqun

APPLICANT: Wang, Aijun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15

CURRENT APPLICATION NUMBER: US/09/736,457

CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 233

LENGTH: 600

TYPE: DNA

ORGANISM: Homo sapien

US-09-736-457-233

Query Match 10.6%; Score 121; DB 4; Length 600;

Best Local Similarity 100.0%; Pred. No. 3.7e-53;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1023 AAGTATATTGTAGAAAGCACTCCCTCAGTCACGCTGTTTAGAGCTGACCAACCAATT 1082

Db 600 AAGTATATTGTAGAAAGCACTCCCTCAGTCACGCTGTTTAGAGCTGACCAACCAATT 541

Qy 1083 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142

Db 540 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481

Qy 1143 A 1143

Db 480 A 480

RESULT 8

US-08-472-659-32

Sequence 32, Application US/08472659

Patent No. 5831030

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5831030uc

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5831030uhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAUCHI, Kozo

APPLICANT: YAMAGUCHI, No. 5831030omi

COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/474,661  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-067339  
FILING DATE: 04-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: REA, TERESA STANEK  
REGISTRATION NUMBER: 30,427  
REFERENCE/DOCKET NUMBER: 001560-204  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-6620  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-474-661-32

Query Match 2.2%; Score 25; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTCCCTGCTGCAGCAATG 25  
Db 13 ATGGCTCCCTGCTGCAGCAATG 37

RESULT 10  
US-08-611-977-32  
Sequence 32, Application US/08611977  
Patent No. 5972886  
GENERAL INFORMATION:  
APPLICANT: TSUJIMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSUROUKA, No. 5972886u  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. 5972886uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAUCHI, Kozo  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,661  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-067339  
FILING DATE: 04-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-204  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-6620  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-611-977-32

Query Match 2.2%; Score 25; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTCCCTGCTGCAGCAATG 25  
Db 13 ATGGCTCCCTGCTGCAGCAATG 37

RESULT 11  
US-08-472-659-14  
Sequence 14, Application US/08472659  
Patent No. 5831030  
GENERAL INFORMATION:  
APPLICANT: TSUJIMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSUROUKA, No. 5831030u  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. 5831030uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAUCHI, Kozo  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,659  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
PRIOR APPLICATION DATA:

us-09-508-997a-1.rni

Mon Dec 15 08:51:17 2003

```

; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; INFORMATION FOR SEQ ID NO: 14:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-472-659-14

```

```

Query Match 2.0%; Score 23; DB 2; Length 23;
Best Local Similarity 100.0%; Pred.No. 0.034;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 791 CAAGCGGAATGACCTCTAAGTAT 813
DB 1 CAAGCGGAATGACCTCTAAGTAT 23

```

```

RESULT 12
US-08-474-661-14
; Sequence 14, Application US/08/474661
; Patent No. 5874253
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. 5874253uo
; APPLICANT: NAKAZATO, Hiroschi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5874253uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,661
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; INFORMATION FOR SEQ ID NO: 14:
; NAME: REA, TERESA STANEX
; REGISTRATION NUMBER: 30,427

```

```

; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-474-661-14

```

```

Query Match 2.0%; Score 23; DB 2; Length 23;
Best Local Similarity 100.0%; Pred.No. 0.034;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 791 CAAGCGGAATGACCTCTAAGTAT 813
DB 1 CAAGCGGAATGACCTCTAAGTAT 23

```

```

RESULT 13
US-08-611-977-14
; Sequence 14, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroschi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; INFORMATION FOR SEQ ID NO: 14:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs

```

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-611-977-14

Query Match 2.0%; Score 23; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.034;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 CAAGCGGAATGACCTCTAAGTAT 813  
DB 1 CAAGCGGAATGACCTCTAAGTAT 23

RESULT 14  
US-08-472-659-15/c  
Sequence 15, Application US/08472659  
Patent No. 5831030  
GENERAL INFORMATION:  
APPLICANT: TSUTSUMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSUROOKA, No. 5831030uo  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. 5831030uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAICHI, Kozo  
APPLICANT: YAMAGUCHI, No. 5831030omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,659  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-067339  
FILING DATE: 04-MAR-1993  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-248  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-472-659-15

Query Match 1.8%; Score 21; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 CTCCTCTGGGATTGCTTCGGGG 939  
DB 21 CTCCTCTGGGATTGCTTCGGGG 1

RESULT 15  
US-08-472-659-16/c  
Sequence 16, Application US/08472659  
Patent No. 5831030  
GENERAL INFORMATION:  
APPLICANT: TSUTSUMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSUROOKA, No. 5831030uo  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. 5831030uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAICHI, Kozo  
APPLICANT: YAMAGUCHI, No. 5831030omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,659  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-067339  
FILING DATE: 04-MAR-1993  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-248  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-472-659-16

Query Match 1.8%; Score 21; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 709 ATGTACGTTCTGCTGCTGAG 729  
DB 21 ATGTACGTTCTGCTGCTGAG 1

us-09-508-997a-1.rni

Mon Dec 15 08:51:17 2003

Search completed: December 14, 2003, 03:21:09  
Job time : 101 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2003, 03:21:22 ; Search time 4369 Seconds  
(without alignments)  
10702.610 Million cell updates/sec

Title: US-09-508-997a-1

Perfect score: 1143  
Sequence: 1 atggcctcccttgcgtgcagc.....gcaagtttttgccttga 1143

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pia.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query		ID	Description
		Match	Length		
1	1143	100.0	1143	6	BD094240 Model ani
2	1143	100.0	1950	6	AR052445 Sequence
3	1143	100.0	1950	6	AR082433 Sequence
4	1143	100.0	1950	6	E08396 cDNA encodi
5	1143	100.0	1950	9	D88575 Homo sapien
6	1143	100.0	2249	9	AF027866 Homo sapi
C 7	399	34.9	157284	9	AC072051 Homo sapi
8	399	34.9	189092	2	AC072524 Homo sapi
9	254	22.2	254	6	AR270095 Sequence
10	230	20.1	316	6	AR270261 Sequence
11	121	10.6	600	6	AR272521 Sequence
C 12	121	10.6	600	6	AR276102 Sequence
C 13	121	10.6	600	6	AX062606 Sequence
C 14	121	10.6	600	6	AX367523 Sequence
15	31	2.7	1143	10	AF105329 Rattus no
16	31	2.7	1147	6	BD094242 Model ani
17	31	2.7	1229	6	BD094241 Model ani
18	31	2.7	1229	6	BD096261 Method fo
19	31	2.7	1386	10	AF105328 Mus muscu
20	31	2.7	121141	10	AC125314 Mus muscu
21	30	2.6	30	6	BD094249 Model ani
22	30	2.6	44	6	BD094243 Model ani
23	29	2.5	66599	2	AC129921 Mus muscu
24	29	2.5	199669	10	AC103453 Rattus no
C 25	29	2.5	219071	2	AC133259 Rattus no
C 26	26	2.3	26	6	BD094244 Model ani
27	25	2.2	37	6	AR052447 Sequence
28	25	2.2	37	6	AR082435 Sequence
29	25	2.2	37	6	E08397 PCR primer
30	23	2.0	23	6	AR052429 Sequence
31	23	2.0	23	6	AR082417 Sequence
32	23	2.0	142127	10	AC122521 Mus muscu
C 33	23	2.0	170401	2	BX294444 Danio rer
34	23	2.0	198515	10	AF129005 Mus muscu
C 35	23	2.0	285727	2	AC025117 Mus muscu
36	22	1.9	19950	3	AF267746 Drosophil
C 37	22	1.9	51602	2	AC129261 Rattus no
C 38	22	1.9	89248	2	AC021919 Homo sapi
C 39	22	1.9	144430	9	AC091155 Homo sapi
40	22	1.9	168509	9	AC009837 Homo sapi
C 41	22	1.9	171627	2	AC126924 Sus scrof
C 42	22	1.9	172195	9	AC093862 Homo sapi
C 43	21	1.8	21	6	AR052430 Sequence
C 44	21	1.8	21	6	AR052431 Sequence
C 45	21	1.8	21	6	AR052442 Sequence

## ALIGNMENTS

RESULT 1  
BD094240 1143 bp DNA linear PAT 27-AUG-2002  
LOCUS Model animals of nephropathy proliferating mesangial cells.  
DEFINITION BD094240  
ACCESSION BD094240  
VERSION BD094240.1 GI:22639828  
KEYWORDS WO 0124628-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1143)  
AUTHORS Miyata,T.  
TITLE Model animals of nephropathy proliferating mesangial cells  
JOURNAL Patent: WO 0124628-A 1 12-APR-2001;

Mon Dec 15 08:51:16 2003

TOSHIO MIYATA, KIYOSHI KUROKAWA  
 OS Homo sapiens (human)  
 FN WO 0124628-A/1  
 PD 12-APR-2001  
 PF 06-OCT-2000 WO 2000JP006988  
 PR 06-OCT-1999 JP 99P 285736  
 PI TOSHIO MIYATA  
 PC A01K67/027,A61P13/12,A61K45/00,C12N15/12,C12N15/85,G01N33/15,  
 CC G01N33/50  
 PH Key Location/Qualifiers  
 FT CDS Location/Qualifiers  
 FEATURES  
 source  
 i..1143  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 336 a 230 c 261 g 316 t  
 BASE COUNT  
 ORIGIN  
 Query Match 100.0%; Score 1143; DB 6; Length 1143;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGCTCCCTTGTGCGAGAAATGCGAAATGCGAGATTTTGTCTCAACCTGTTGAGAGATGATGAT 60  
 DB 1 ATGGCTCCCTTGTGCGAGAAATGCGAAATGCGAGATTTTGTCTCAACCTGTTGAGAGATGATGAT 60  
 QY 61 GACAAATCAAGGAATGGAATGTTCTTTTCTCTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 DB 61 GACAAATCAAGGAATGGAATGTTCTTTTCTCTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 QY 121 CTGCTCGCT 180  
 DB 121 CTGCTCGCT 180  
 QY 181 AACACTGCTCAGATATGGAATGTTCTTTCTAATAGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 240  
 DB 181 AACACTGCTCAGATATGGAATGTTCTTTCTAATAGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 240  
 QY 241 AAAAGAGTTTCTGATATAAATGCTCCCAAGATATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 241 AAAAGAGTTTCTGATATAAATGCTCCCAAGATATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 301 GGGCTTTTCTGAAAGCTATGCTGCTTTCTAATAGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 360  
 DB 301 GGGCTTTTCTGAAAGCTATGCTGCTTTCTAATAGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 360  
 QY 361 TTATACGATGCCAAGTGGAGCGAGTTGACTTTTACGAATCTATTAGAGACACTAGAGCT 420  
 DB 361 TTATACGATGCCAAGTGGAGCGAGTTGACTTTTACGAATCTATTAGAGACACTAGAGCT 420  
 QY 421 AATATAATAGTGGGTTGAAATGAACACATGCGCAAAATCAAGACGATGATGCTGAA 480  
 DB 421 AATATAATAGTGGGTTGAAATGAACACATGCGCAAAATCAAGACGATGATGCTGAA 480  
 QY 481 GGTGGCATAGCTCATCTGCTGCTAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 DB 481 GGTGGCATAGCTCATCTGCTGCTAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 QY 541 TGGCATAGCTTCCACCAAGAGCGAATCCATTAATGTCATTTCAATCTCCCAAGTGC 600  
 DB 541 TGGCATAGCTTCCACCAAGAGCGAATCCATTAATGTCATTTCAATCTCCCAAGTGC 600  
 QY 601 TGTGGAGGAGCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 DB 601 TGTGGAGGAGCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 QY 661 GACCCATCAATGAGATTTCTGAGCTTCAGATACATGATGATGATGATGATGATGATGATGATGATGAT 720  
 DB 661 GACCCATCAATGAGATTTCTGAGCTTCAGATACATGATGATGATGATGATGATGATGATGATGATGAT 720  
 QY 721 CTGCTGAGAT 780

DB 721 CTGCTGAGAT 780  
 QY 781 TGGCAATCCCAAGGCGAATGACCTCTAAGTATGTTGAGTATGTTTCTCTGAGTTCAAG 840  
 DB 781 TGGCAATCCCAAGGCGAATGACCTCTAAGTATGTTGAGTATGTTTCTCTGAGTTCAAG 840  
 QY 841 ATAGAGAGATTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 900  
 DB 841 ATAGAGAGATTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 900  
 QY 901 GATGATCCCAAGCAGATCTCTGCGGATGCTTCCGGGGGGTCTGCTCTCTATATATCAAGG 960  
 DB 901 GATGATCCCAAGCAGATCTCTCTGGGATGCTTCCGGGGGGTCTGCTCTCTATATATCAAGG 960  
 QY 961 ATGATGCACAAAATCTTACATAGAGTCTAGAGAGGCGACCGAGGCTACTGCTGCGCAC 1020  
 DB 961 ATGATGCACAAAATCTTACATAGAGTCTAGAGAGGCGACCGAGGCTACTGCTGCGCAC 1020  
 QY 1021 GGAAGTAAATTTGTAAGAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080  
 DB 1021 GGAAGTAAATTTGTAAGAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080  
 QY 1081 TTCTCTATTTGTTATCAGGAAGATGACATCATCTTATTCAGTGGCAAGTTCTTGGCCT 1140  
 DB 1081 TTCTCTATTTGTTATCAGGAAGATGACATCATCTTATTCAGTGGCAAGTTCTTGGCCT 1140  
 QY 1141 TGA 1143  
 DB 1141 TGA 1143  
 RESULT 2  
 AR052445 1950 bp DNA linear PAT 29-SEP-1999  
 LOCUS Sequence 30 from patent US 5831030.  
 DEFINITION AR052445  
 ACCESSION AR052445  
 VERSION AR052445.1 GI:5975809  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 1950)  
 AUTHORS Teiimoto,M., Iwasa,F., Teiuroka,N., Nakazato,H., Miura,K.,  
 Ishida,N., Kurihara,T., Yamachi,K. and Yamaguchi,N.  
 TITLE Antibodies specific for megakaryocyte differentiation factor  
 JOURNAL Patent: US 5831030-A 30 03-NOV-1998;  
 FEATURES Location/Qualifiers  
 source 1..1950  
 BASE COUNT 570 a 384 c 407 g 589 t  
 ORIGIN  
 Query Match 100.0%; Score 1143; DB 6; Length 1950;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGCTCCCTTGTGCGAGAAATGCGAAATGCGAGATTTTGTCTCAACCTGTTGAGAGATGATGAT 60  
 DB 74 ATGGCTCCCTTGTGCGAGAAATGCGAAATGCGAGATTTTGTCTCAACCTGTTGAGAGATGATGAT 133  
 QY 61 GACAAATCAAGGAATGGAATGTTCTTTTCTCTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 DB 134 GACAAATCAAGGAATGGAATGTTCTTTTCTCTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 193  
 QY 121 CTGCTCGCT 180  
 DB 134 CTGCTCGCT 253  
 QY 181 AACACTGCTCAGATATGGAATGTTCTTTCTAATAGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 240  
 DB 254 AACACTGCTCAGATATGGAATGTTCTTTCTAATAGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 313



ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1950)
TITLE	Tsujiimoto,M., Iwasa,F., Tsuruoka,N., Nakazato,H., Miura,K., Ishida,N., Kurihara,T., Yamaichi,K. and Yamaguchi,N.
JOURNAL	Megakaryocyte differentiation factor
FEATURES	Patent: US 5972886-A 30 26-OCT-1999; Location/Qualifiers 1..1950 /organism="unknown"
BASE COUNT	570 a 384 c 407 g 589 t
ORIGIN	
Query Match	100.0%; Score 1143; DB 6; Length 1950;
Best Local Similarity	100.0%; Pred No. 0;
Matches 1143; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGGCCTCCCTTGCCTGCACAAATGCAGAGTTTCTTCAACCTGTTCAGAGAGATCGAT 60
Db	74 ATGGCCTCCCTTGCCTGCACAAATGCAGAGTTTCTTCAACCTGTTCAGAGAGATCGAT 133
Qy	61 GACATCAAGGAATAGGAATGTGTTCTTCTCTCTGAGCCTCTTCGCTGCCCTGGCC 120
Db	134 GACATCAAGGAATAGGAATGTGTTCTTCTCTCTGAGCCTCTTCGCTGCCCTGGCC 193
Qy	121 CTGGTCGGCTTGGCGCTCAAGATGACTCCCTCTCTCAGATTCATGAATGATGCTCATGTT 180
Db	194 CTGGTCGGCTTGGCGCTCAAGATGACTCCCTCTCTCAGATTCATGAATGATGCTCATGTT 253
Qy	181 AACACTGGCTCAGGATATGGAAACTCTTCTTAATGATCAGTCAGGCTCCAGTCTCAACTG 240
Db	254 AACACTGGCTCAGGATATGGAAACTCTTCTTAATGATCAGTCAGGCTCCAGTCTCAACTG 313
Qy	241 AAAAGAGTTTTTCTGATATAAATGCATCCCAAGGATATGATCTCAGCAATGTGCAAT 300
Db	314 AAAAGAGTTTTTCTGATATAAATGCATCCCAAGGATATGATCTCAGCAATGTGCAAT 373
Qy	301 GGGCTTTTGTGTAAGAAAGTATGGCTTTTCATAGGACTACATTCAGTGTGCCGAAAAA 360
Db	374 GGGCTTTTGTGTAAGAAAGTATGGCTTTTCATAGGACTACATTCAGTGTGCCGAAAAA 433
Qy	361 TTATACGATGCCAAGTGGAGCGAGTTGACTTTACGATCATTTAGAGACATAGACGT 420
Db	434 TTATACGATGCCAAGTGGAGCGAGTTGACTTTACGATCATTTAGAGACATAGACGT 493
Qy	421 AATATTAATAGTGGTGTGAAATCAACACATGGCAATCAAGACGTGATGGTGAA 480
Db	494 AATATTAATAGTGGTGTGAAATCAACACATGGCAATCAAGACGTGATGGTGAA 553
Qy	481 GGTGGCATTAAGCTCATCTGCTGTGAATGGTGTGCTGAATGCTGTACTTCAAAGCGAAG 540
Db	554 GGTGGCATTAAGCTCATCTGCTGTGAATGGTGTGCTGAATGCTGTACTTCAAAGCGAAG 613
Qy	541 TGGCAATCAGCCTTCCACAGAGCGAAACCATTAATGGCAATTCAAATCTCCAGTGC 600
Db	614 TGGCAATCAGCCTTCCACAGAGCGAAACCATTAATGGCAATTCAAATCTCCAGTGC 673
Qy	601 TCTGGGAAGGAGTGGCCATGATCATCAGAGCGGAAGTTCAATTTGCTGTATTTCAG 660
Db	674 TCTGGGAAGGAGTGGCCATGATCATCAGAGCGGAAGTTCAATTTGCTGTATTTCAG 733
Qy	661 GACCCATCAATGAAGATCTTGAGCTCAGATACAAATGGTGGCATAAACATGTACGTTCTG 720
Db	734 GACCCATCAATGAAGATCTTGAGCTCAGATACAAATGGTGGCATAAACATGTACGTTCTG 793
Qy	721 CTGCTTGAGATGACCTCTCTGAAATCGAAAAAATCTGACCTTTCAGATCTAATGGAA 780
Db	794 CTGCTTGAGATGACCTCTCTGAAATCGAAAAAATCTGACCTTTCAGATCTAATGGAA 853
Qy	781 TGGACCAATCCAGCGCAATGACCTTAAGTATGTTGAGTATTTTTTTCCTCAGTTTCAAG 840
Db	854 TGGACCAATCCAGCGCAATGACCTTAAGTATGTTGAGTATTTTTTTCCTCAGTTTCAAG 913

Qy	721	CTGCTGAGAAAGACCTCTCTGAAATCGAAACAAACCTGACCTTCAGAAATCTAAAGAA	785
Db	794	CTGCTGAGAAAGACCTCTCTGAAATCGAAACAAACCTGACCTTCAGAAATCTAAAGAA	853
Qy	781	TGGACCAATCCAAAGCGGAATGACCTCTAAGATATGTTTGAAGTATTTTTTCTCTAGTCTCAAG	840
Db	854	TGGACCAATCCAAAGCGGAATGACCTCTAAGATATGTTTGAAGTATTTTTTCTCTAGTCTCAAG	913

```

Query Match      100.0%; Score 1143; DB 6; Length 1950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BASE COUNT      570 a      384 c      407 g      589 t
ORIGIN

1  ATGCGCTCCCTTCGTCGAGAGAAATGAGAGTTTTCCTTCAACCTGTCAGAGAGATGGAT 60
Dbb
74  ATGCGCTCCCTTCGTCGAGAGAAATGAGAGTTTTCCTTCAACCTGTCAGAGAGATGGAT 133
Qy
61  GACATCAAGGAAATGGAATGTTTCTTCTCTGAGAGCTTTCGCTGCGCTGCGC 120
Dbb
134  GACAAATCAAGGAAATGGAATGTTTCTTCTCTGAGAGCTTTCGCTGCGCTGCGC 193
Qy
121  CTGCTGCGCTTGGCGCTCAAGATGACTCCCTCTCTCAGATGATTAAGTTGCTTCATGTT 180
Dbb
194  CTGCTGCGCTTGGCGCTCAAGATGACTCCCTCTCTCAGATGATTAAGTTGCTTCATGTT 253
Qy
181  AACACTGCTCAGGATATGGAATCTCTCTTAATAGTCAAGTCCGCTCCAGTCTCACTG 240
Dbb
254  AACACTGCTCAGGATATGGAATCTCTCTTAATAGTCAAGTCCGCTCCAGTCTCACTG 313
Qy
241  AAAGAGTCTTCTGATATTAATGATCCCAAGGATTAATGATCTCAGCATTTGTAAT 300
Dbb
314  AAAGAGTCTTCTGATATTAATGATCCCAAGGATTAATGATCTCAGCATTTGTAAT 373
Qy
301  GGGCTTTTGTCTGAAAGAGTGTATGGCTTCTAAGGCTACATTCAGTGTGCGGAAAA 360
Dbb
374  GGGCTTTTGTCTGAAAGAGTGTATGGCTTCTAAGGCTACATTCAGTGTGCGGAAAA 433
Qy
361  TTATACGATCCCAAGAGTGTATGGCTTCTAAGGCTACATTCAGTGTGCGGAAAA 420
Dbb
434  TTATACGATCCCAAGAGTGTATGGCTTCTAAGGCTACATTCAGTGTGCGGAAAA 493
Qy
421  AATATTAATTAAGTGGGTGAAATGAAACACATGCGCAAAATCAAGACGTGATGTA 480
Dbb
494  AATATTAATTAAGTGGGTGAAATGAAACACATGCGCAAAATCAAGACGTGATGTA 553
Qy
481  GGTGGCATAAGCTCATCTGCTGTAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 540
Dbb
554  GGTGGCATAAGCTCATCTGCTGTAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 613
Qy
541  TGGCATAAGCTCATCTGCTGTAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 600
Dbb
614  TGGCATAAGCTCATCTGCTGTAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 673
Qy
601  TCTGGGAGGCGAGTGGCCTGATGATGATGATGATGATGATGATGATGATGATGATG 660
Dbb
674  TCTGGGAGGCGAGTGGCCTGATGATGATGATGATGATGATGATGATGATGATGATG 733
Qy
661  GACCCATCAATCAAGATTTCTGAGCTCAGATGATGATGATGATGATGATGATGATG 720
Dbb
734  GACCCATCAATCAAGATTTCTGAGCTCAGATGATGATGATGATGATGATGATGATG 793
Qy
721  CTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Dbb
794  CTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 853
Qy
781  TGGACCAATCCAGGCGAATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Dbb
854  TGGACCAATCCAGGCGAATGATGATGATGATGATGATGATGATGATGATGATGATG 913
Qy
841  ATAGAGAGAGATTTATGAAATGAAACCAATATTTGAGAGCCCTAGGCTGGAAGATAT 900
Dbb
914  ATAGAGAGAGATTTATGAAATGAAACCAATATTTGAGAGCCCTAGGCTGGAAGATAT 973
Qy
901  GATGATCCAAAGCAGATCTCTCTGAGGATGCTTCTGGGGGGTCTCTGTATATATCAAG 960
Dbb
974  GATGATCCAAAGCAGATCTCTCTGAGGATGCTTCTGGGGGGTCTCTGTATATATCAAG 1033

```

```

/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
570 a      384 c      407 g      589 t

```

RESULT 4

```

E08396      1950 bp      RNA      linear      PAT 29-SEP-1997
LOCUS      CDNA encoding human megakaryocyte growth differentiating factor.
DEFINITION      E08396
ACCESSION      E08396
VERSION      E08396.1 GI:2176513
KEYWORDS      JP 1994313000-A/20.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Tsuimoto,M., Kurihara,T., Ishida,N., Iwasa,F., Nakazato,H.,
Yamaichi,H., Miura,T., Tsuruoka,N. and Yamaguchi,M.
MEGAKARYOCYTE-PROLIFERATING AND DIFFERENTIATING FACTOR
Patent: JP 1994313000-A 20 08-NOV-1994;
JOURNAL      SUNTORY LTD
COMMENT      OS Homo sapiens (human)
PN JP 1994313000-A/20
PD 08-NOV-1994
PF 16-JUL-1993 JP 1993197752
PR 17-JUL-1992 JP 92P 212205, 04-MAR-1993 JP 93P 67339 PI
TSUJIMOTO MASAFUMI, KURIHARA TATSUYA, ISHIDA NOBUHIRO, PI IWASA
FUYUKI,
PI NAKAZATO HIROSHI, YAMAICHI HIROZO, MIURA TAKEHISA, PI
TSURUOKA NOBUO,
PI YAMAGUCHI WARE
PC C07K15/14,A61K37/02,C12N5/10,C12N15/19,C12P21/02,(C12P21/02,
CC C12R1:191);
CC Strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..1950
FH /organism="Homo sapiens"
FH /cell_line="A431 (human epidermoid carcinoma)"
FH 5'UTR 1..73
FH CDS 74..1216
FH /product="Human megakaryocyte growth factor"
FH differentiating factor
FH 3'UTR 1217..1950
FH polyA signal 1933..1938.
FH Location/Qualifiers
FH 1..1950

```

961 ATGATGCAAAATCTTACATAGAGGTCACTGAGGAGGCAACGAGGCTACTGTCGCACA 1020  
1034 ATGATGCAAAATCTTACATAGAGGTCACTGAGGAGGCAACGAGGCTACTGTCGCACA 1093  
1021 GGAAGTAATATGTAGAAAAGCAATCCCTCAGTCACTGAGGCTACTGTCGCACA 1080  
1094 GGAAGTAATATGTAGAAAAGCAATCCCTCAGTCACTGAGGCTACTGTCGCACA 1153  
1081 TTCCATATTTGTTATCAGAGGATGACATCATCTTATTCACTGGCAAGTTCTTGCCT 1140  
1154 TTCCATATTTGTTATCAGAGGATGACATCATCTTATTCACTGGCAAGTTCTTGCCT 1213  
1141 TGA 1143  
1214 TGA 1216  
RESULT 5  
D88575 1950 bp mRNA linear PRI 04-JUL-1998  
LOCUS Homo sapiens mRNA for TP55, complete cds.  
DEFINITION D88575  
ACCESSION D88575.1 GI:3288674  
VERSION TP55.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Tsujimoto, M., Tsuruoka, N., Ishida, N., Kurihara, T., Iwasa, P.,  
Yamashiro, K., Rogi, T., Kodama, S., Katsuragi, N., Adachi, M.,  
Kawashima, T., Nakao, M., Yamauchi, K., Hashino, J., Haruyama, M.,  
Miura, K., Nakanishi, T., Nakazato, H., Teramura, M., Mizoguchi, H. and  
Yamaguchi, N.  
Purification, cDNA cloning, and characterization of a new serpin  
with megakaryocyte maturation activity  
J. Biol. Chem. 272 (24), 15373-15380 (1997)  
MEDLINE 97326116  
PUBMED 9182567  
REFERENCE 2 (bases 1 to 1950)  
Tsuruoka, N.  
Direct Submission  
Submitted (25-OCT-1996) Nobuo Tsuruoka, Sunitory Institute for  
Biomedical Research, Pharmaceutical Research Laboratories I; 1-1-1  
Wakayamadai, Shinamoto-cho, Mishima-gun, Osaka 618, Japan  
(E-mail:tsuruoka@minase.sunitory.co.jp, Tel:075-962-9285 (ex.3110),  
Fax:075-962-6448)  
FEATURES  
Location/Qualifiers  
1..1950  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="A431"  
/cell\_type="epidermoid carcinoma"  
74..1216  
/codon\_start=1  
/product="tp55"  
/protein\_id="BAA31232.1"  
/db\_xref="GI:3288675"  
CDS  
/translation="MASHLAANAIFCFNLFRMDNQGNGVYFSSLSLFAALVRL  
GAQDLSQIDKLHVNTASGNSQSGHQSLKRVFSDINASHKADYDLSYNGL  
FAKVTGPHVYIECAKLYDAKVERVFTNHLEDTRNKVENHFKIKNVIGE  
GGTSSAVMLVNAVYFKGWSQAFTKSETTNCHEFKSPKSGKAVAMMHQERKENS  
LEDESKVLLIYNGGINMYVLLPNDLSEIKNLTFONLMEWTNPRMTSKYVEVFF  
PQYKIKENTMKQYLRALGKIDIPESKADLSGIASGRLYISRMVHKSLEYTEVETG  
EATAATGSMVNEKQLPQSTFLRDPHFVLFVIRKDDILFSGKVSCEP"  
BASE COUNT 570 a 384 c 407 g 589 t  
ORIGIN  
Query Match 100.0%; Score 1143; DB 9; Length 1950;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGTCTGCGCAAAATGCAGAGTTTGTTCAACTCTTCAGAGAGATGGAT 60  
DB 74 ATGGCTCCCTTGTCTGCGCAAAATGCAGAGTTTGTTCAACTCTTCAGAGAGATGGAT 133  
QY 61 GACAAATCAAGAGAAATGGAAATGTGTCTCTCTGAGGCTCTTCGCTGCGCTGGCC 120  
DB 134 GACAAATCAAGAGAAATGGAAATGTGTCTCTCTGAGGCTCTTCGCTGCGCTGGCC 193  
QY 121 CTGGTCCGCTTGGGGCGCTCAAGATGATGATCCCTCTCTCAGATGATGATGATGATGAT 180  
DB 194 CTGGTCCGCTTGGGGCGCTCAAGATGATGATCCCTCTCTCAGATGATGATGATGATGAT 253  
QY 181 AACACTCGCTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
DB 254 AACACTCGCTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 313  
QY 241 AAAAGAGTTTCTGATATAAATGCATCCCAAGGATTAATGATCCTCAGCAATGTGAAT 300  
DB 314 AAAAGAGTTTCTGATATAAATGCATCCCAAGGATTAATGATCCTCAGCAATGTGAAT 373  
QY 301 GGGCTTTTCTGTAAGAAATGTATGGCTTTCAATAGGACTACATTCAGTGTGCGGAAAA 360  
DB 374 GGGCTTTTCTGTAAGAAATGTATGGCTTTCAATAGGACTACATTCAGTGTGCGGAAAA 433  
QY 361 TTATACGATGCCAAAGTGGAGCGTGTGATTTTACGAATCAATTTAGAGACACATAGACGT 420  
DB 434 TTATACGATGCCAAAGTGGAGCGTGTGATTTTACGAATCAATTTAGAGACACATAGACGT 493  
QY 421 AATATTAAATAGTGGGTGAAATGAAACACATGCAAAATCAAGAACTGTATGGTGA 480  
DB 494 AATATTAAATAGTGGGTGAAATGAAACACATGCAAAATCAAGAACTGTATGGTGA 553  
QY 481 GTGGCATTAAGTCAATCTGCTGTAAGTGTGCTGTAAGTGTGCTGTAAGTGTGCTGTAAG 540  
DB 554 GTGGCATTAAGTCAATCTGCTGTAAGTGTGCTGTAAGTGTGCTGTAAGTGTGCTGTAAG 613  
QY 541 TGGCAATTCAGCTTCCCAAGAGGCAAAACCAATAATTCGCAATTCGCAATTCGCAATTCG 600  
DB 614 TGGCAATTCAGCTTCCCAAGAGGCAAAACCAATAATTCGCAATTCGCAATTCGCAATTCG 673  
QY 601 TCTGGAGGAGGAGTCCCATGATGATCAGGAAACGAAAGTTCAATTTGTCTGTTATTGAG 660  
DB 674 TCTGGAGGAGGAGTCCCATGATGATCAGGAAACGAAAGTTCAATTTGTCTGTTATTGAG 733  
QY 661 GACCATCAATGAGAAATCTTGAGCTCAGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
DB 734 GACCATCAATGAGAAATCTTGAGCTCAGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793  
QY 721 CTGGCTGAGATGACCTCTCTGAAATGAAACCAAACTGACCTTTCAGAAATCTTAATGGAA 780  
DB 794 CTGGCTGAGATGACCTCTCTGAAATGAAACCAAACTGACCTTTCAGAAATCTTAATGGAA 853  
QY 781 TGGACCAATCCAAAGGCGAATGACCTCTGAGTATGTTGAGGTTATTTTTCCTCAGTTCAG 840  
DB 854 TGGACCAATCCAAAGGCGAATGACCTCTGAGTATGTTGAGGTTATTTTTCCTCAGTTCAG 913  
QY 841 ATAGAGAAATATTAATGAAATGAAACCAATATTTTCAGAGCCCTAGGCTGAAAGATATCTTT 900  
DB 914 ATAGAGAAATATTAATGAAATGAAACCAATATTTTCAGAGCCCTAGGCTGAAAGATATCTTT 973  
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGTTCGCTGTATATATCAAGG 960  
DB 974 GATGAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGTTCGCTGTATATATCAAGG 1033  
QY 961 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGCGCACCGAGGCTACTGCTGCCACA 1020  
DB 1034 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGCGCACCGAGGCTACTGCTGCCACA 1093  
QY 1021 GGAAGTAATATGTAGAAAAGCAATCCCTCAGTCACTGAGGCTGTTTAGAGCTGACCAACCA 1080  
DB 1094 GGAAGTAATATGTAGAAAAGCAATCCCTCAGTCACTGAGGCTGTTTAGAGCTGACCAACCA 1153

Mon Dec 15 08:51:16 2003

```

1081 TTCCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1140
Db 1154 TTCCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1213
QY 1141 TGA 1143
Db 1214 TGA 1216

RESULT 6
AF027866 2249 bp mRNA linear PRI 20-OCT-1998
LOCUS
DEFINITION Homo sapiens megin mRNA, complete cds.
ACCESSION AF027866
VERSION AF027866.1 GI:3769372
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Miyata, T., Nangaku, M., Suzuki, D., Inagi, R., Uragami, K., Sakai, H., Okubo, K. and Kurokawa, K.
TITLE A mesangium-predominant gene, megin, is a new serpin upregulated in IGA nephropathy
JOURNAL J. Clin. Invest. 102 (4), 828-836 (1998)
MEDLINE 98376492
PUBMED 9710452
REFERENCE
AUTHORS Toshio, M., Yasuda, Y., Nangaku, M. and Kurokawa, K.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-1997) Medical Science, Tokai University, Bousaidai, Isehara, Kanagawa 259-11, Japan
3 (bases 1 to 2249)
REFERENCE
AUTHORS Toshio, M., Yasuda, Y., Nangaku, M. and Kurokawa, K.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1998) Medical Science, Tokai University, Bousaidai, Isehara, Kanagawa 259-11, Japan
Sequence update by submitter
On Oct 20, 1998 this sequence version replaced gi:3764058.
FEATURES
source
1..2249
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="mesangial"
365..1507
/codon_start=1
/product="megin"
/protein_id="AAC64506.1"
/db_xref="GI:3769372"
/translation="MASLAAAEAEFCNLPRMDDNGNGVFTSSLSLFAALALVRL
GADPSLSQIDKLHVNATAGYGNSSQSLQSLKRVFSDINASHKDYDLSIVNGL
FASKVYFKDYIEAEKLYDAKVERVDFTHLEDTRNINWVNEHGHKIKNVIG
GGTSSSAWMLVNAVYFKGWQSAFTKSETINHFSPKSCSAVAMVHQRKELSV
IEDPSMKLLELVNGGINMVLIPNDLSIEINKLTFQNLMEWTPRMTSKIVEVFF
PQKIEKNEMKQVIRALGLDIFDSKADLSGIASGRLYISRMWHSYIRVTEBGT
EATAAGSNIVEXPLQFSTLFRADHPFLVIRKDDILFSGKVSQP"
BASE COUNT 649 a 452 c 461 g 683 t 4 others
ORIGIN
Query Match 100.0%; Score 1143; DB 9; Length 2249;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTCCCTTGCTGCAGCAATGCAGATTTGCTTCAACCTGTTTCAGAGATGAT 424
Db 365 ATGGCTCCCTTGCTGCAGCAATGCAGATTTGCTTCAACCTGTTTCAGAGATGAT 424
QY 61 GACAAATCAGGAATGGAATGTGTTCTTCTCTCTGAGCTCTTGCTGCTGGCCG 120
Db 425 GACAAATCAGGAATGGAATGTGTTCTTCTCTCTGAGCTCTTGCTGCTGGCC 484

```

```

121 CTGGTCCGCTTGGCGCTCAAGTGAATCTCCCTCTCTCAGATTGATAAGTTGCTTCATCTT 180
Db 485 CTGGTCCGCTTGGCGCTCAAGTGAATCTCCCTCTCTCAGATTGATAAGTTGCTTCATCTT 544
QY 181 AACACTGGCTCAGATATGGAACCTCTTCTAAAGTCACTCAGAGGCTCCAGTCTCACTG 240
Db 545 AACACTGGCTCAGATATGGAACCTCTTCTAAAGTCACTCAGAGGCTCCAGTCTCACTG 604
QY 241 AAAGAGTTTCTTCTGATATAATGCATCCCAAGAGTTATGATCTCAGCAATGTGAAT 300
Db 605 AAAGAGTTTCTTCTGATATAATGCATCCCAAGAGTTATGATCTCAGCAATGTGAAT 664
QY 301 GGGCTTTTCTCGAAAAAGTGTATGCTTCTAAAGCACTACATTTAGTGTGCGAAAAA 360
Db 665 GGGCTTTTCTCGAAAAAGTGTATGCTTCTAAAGCACTACATTTAGTGTGCGAAAAA 724
QY 361 TTATACGATGCCAAGTGGAGCGAGTTGACTTTAGCAATCAATTTAGAAGACACTAGAGCT 420
Db 725 TTATACGATGCCAAGTGGAGCGAGTTGACTTTAGCAATCAATTTAGAAGACACTAGAGCT 784
QY 421 AATATTAAATAGTGGGTTGAAAAATGAAAACAATGGCAAAATCAAGAACGTGATTGGTAA 480
Db 785 AATATTAAATAGTGGGTTGAAAAATGAAAACAATGGCAAAATCAAGAACGTGATTGGTAA 844
QY 481 GGTGGCATAAGCTCATCTGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 845 GGTGGCATAAGCTCATCTGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 904
QY 541 TGGCAATCAGCTTCCACCAAGAGGAAACCAATTAATTTGCCATTTCAAAATCCCAAGTGC 600
Db 905 TGGCAATCAGCTTCCACCAAGAGGAAACCAATTAATTTGCCATTTCAAAATCCCAAGTGC 964
QY 601 TCTGGCAAGGCAATGGCCATGATGATCAGGAACGGAAGTTCATTTGCTGTTATTCAG 660
Db 965 TCTGGCAAGGCAATGGCCATGATGATCAGGAACGGAAGTTCATTTGCTGTTATTCAG 1024
QY 661 GACCCATCAATGAAGATTTCTTGAGCTCAGATACAAATGGTGGCAATAACATGTCGTTG 720
Db 1025 GACCCATCAATGAAGATTTCTTGAGCTCAGATACAAATGGTGGCAATAACATGTCGTTG 1084
QY 721 CTGCTCGAATGACCTCTCTGAAATGAAAACAAACCTGACCTTTCAAGATCTAATGGAA 780
Db 1085 CTGCTCGAATGACCTCTCTGAAATGAAAACAAACCTGACCTTTCAAGATCTAATGGAA 1144
QY 781 TGGACCAATCCAAAGCGGAATGACCTCTAAGTATGTTGAGGTATTTTCTTCAGTTCAG 840
Db 1145 TGGACCAATCCAAAGCGGAATGACCTCTAAGTATGTTGAGGTATTTTCTTCAGTTCAG 1204
QY 841 ATAGAGAGAAATTAATGAATGAACAAATATTTGAGAGCCCTAGGCGTGAAGATATCTTT 900
Db 1205 ATAGAGAGAAATTAATGAATGAACAAATATTTGAGAGCCCTAGGCGTGAAGATATCTTT 1264
QY 901 GATGAATCCAAAGCAGATCTCTCTGGATGTTCTCGGGGGTGTCTGTATATATCAAGG 960
Db 1265 GATGAATCCAAAGCAGATCTCTCTGGATGTTCTCGGGGGTGTCTGTATATATCAAGG 1324
QY 961 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGACCCAGAGGTACTGTGCCACA 1020
Db 1325 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGACCCAGAGGTACTGTGCCACA 1384
QY 1021 GGAATGATATTTAGAAAGCAATCCCTCAGTCCAGCTGTTTGAAGTGAACCA 1080
Db 1385 GGAATGATATTTAGAAAGCAATCCCTCAGTCCAGCTGTTTGAAGTGAACCA 1444
QY 1081 TTCCCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1140
Db 1445 TTCCCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1504
QY 1141 TGA 1143
Db 1505 TGA 1507

```

Ginde, S., Goyette, N., Graman, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Huime, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Kells, C., LaRoque, K., Lamarez, R., Landers, T., Lehoczyk, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPhee, N., Melchior, J., Meneses, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rice, B., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zemke, B., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (03-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Nov 18, 2001 this sequence version replaced gi:14277307.  
All repeats were identified using RepeatMasker:  
Smith, A.P. & Green, P. (1996-1997)

T. A.F.A. & Green, P. (1996-1997)  
 Center: Whitehead Institute/ MIT Center for Genome  
 Science  
 Contact: sequence.submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L366  
 Center clone name: 79 D 21

```

-----
Location/Qualifiers
1. .157284
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-79D21"

```

[illegible]

FEATURES  
source

RESULT 7					
AC072051/c	AC072051	157284 bp	DNA	linear	PRI 03-DEC-2001
LOCUS					
DEFINITION	Homo sapiens chromosome 1, clone RP11-79D21, complete sequence.				
ACCESSION	AC072051				
VERSION	AC072051.8 GI:16974280				
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
REFERENCE	1 (bases 1 to 157284)				
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.				
TITLE	Homo sapiens chromosome 1, clone RP11-79D21				

Unpublished

JOURNAL

2 (Dates 1 to 157284)

REFERENCE

AUTHORS

Barrett, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Boukgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choelpal, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Donino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardy, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M., McSwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Veneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stegase-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrelli, A., Travers, M., Trigliozzi, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Yeung, G., Zainoue, I., Zimmer, A., and Zody, M.

**TITLE**  
JOURNAL  
**REFERENCE**  
**AUTHORS**

Direct Submission  
Submitted (07-JUN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 157284)  
Birken, B., Linton, L., Musbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,  
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
Choepli, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooke, P., DeArellano, K., Dowat, K., Diaz, J.S., Dodge, S., Fero, S.,  
Ferrera, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gort, S., Goyette, X., Graham, L., Grand-Pierre, N.,  
Gins, B., Heasford, A., Horton, L., Hulme, M., Iliev, J., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kelly, C., LaRoque, K.,  
Lanares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, N., McKean, P., McKernan, K., McPheeters, R., Meldrum, J.,  
Mearns, L., Milovan, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,  
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodor, J.,  
Topham, K., Traversman, M., Travis, N., Triglio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zaimov, I., Zembek, I., Zimmer, A. and Zody, M.

**TITLE**  
JOURNAL  
DIRECT SUBMISSION  
Submitted (25-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**REFERENCE**  
4 (bases 1 to 157284)  
**AUTHORS**  
Bixen, B., Linton, L., Nusbaum C., Lander E., Ali, A., Allen, N., Anderson, S., Bara, N., Bastien, V., Boguslavsky, L., Bonkhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Chospel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHuch, W., Gage, D., Galagan, J., Gargyala, S.,

Query Match	Best Local Similarity	100.0%	Pred.No. 1.2e-217	Indels	0	Gaps	0
Matches	399	Conservative	0	Mismatches	0		
repeat_region	complement(14942..15044)						
repeat_region	/rpt_family="MIR3"						
repeat_region	15156..15298						
repeat_region	/rpt_family="L2"						
repeat_region	15592..15773						
repeat_region	/rpt_family="L1M4c"						
repeat_region	15874..18310						
repeat_region	/rpt_family="L1P16"						
unsure	16373..16374						
repeat_region	/note=1128 bp of bacterial transposon insertion in BAC excised between these 2 bp"						
repeat_region	18311..18331						
repeat_region	/rpt_family="(TAAA)n"						
repeat_region	18332..19566						
repeat_region	/rpt_family="L1P16"						
repeat_region	19567..19587						
repeat_region	/rpt_family="(CA)n"						
repeat_region	19588..20042						
repeat_region	/rpt_family="L1P16"						
repeat_region	20043..20064						
repeat_region	/rpt_family="AT rich"						
repeat_region	complement(21259..21386)						
repeat_region	/rpt_family="MIR3"						
repeat_region	22305..22420						
repeat_region	/rpt_family="AT rich"						
repeat_region	complement(22788..22948)						
repeat_region	/rpt_family="TRAM"						
repeat_region	complement(23843..24253)						
repeat_region	/rpt_family="MSTA"						
repeat_region	complement(24678..25041)						
repeat_region	/rpt_family="MLT1A2"						
repeat_region	25054..25094						
repeat_region	/rpt_family="HAL1"						
repeat_region	25214..25358						
repeat_region	/rpt_family="L1M4c"						
repeat_region	complement(25602..25854)						
repeat_region	/rpt_family="MIR"						
repeat_region	26745..26796						
repeat_region	/rpt_family="AT rich"						
repeat_region	27451..27484						
repeat_region	/rpt_family="(CAAAA)n"						
repeat_region	complement(28183..28219)						
repeat_region	/rpt_family="MIR"						
repeat_region	complement(28249..28325)						
repeat_region	/rpt_family="MIR"						
repeat_region	complement(29372..29459)						
repeat_region	/rpt_family="MIR"						
repeat_region	29539..30003						
repeat_region	/rpt_family="MER66B"						
repeat_region	30596..30624						
repeat_region	/rpt_family="AT rich"						
repeat_region	31418..31829						
repeat_region	/rpt_family="L2"						
repeat_region	31846..31915						
repeat_region	/rpt_family="(TCAA)n"						
repeat_region	complement(32019..32103)						
Query Match	34.9%	Score 399	DB 9	Length 157284			
Best Local Similarity	100.0%	Pred.No. 1.2e-217					
Matches	399	Conservative	0	Mismatches	0	Indels	0
QY	745	ATTGAAACAAACTGACCTTTCAGATCTAATGGAATGGAACCAATCCAGGGCAATGCC	804				
Db	126111	ATTGAAACAAACTGACCTTTCAGATCTAATGGAATGGAACCAATCCAGGGCAATGCC	126052				
QY	805	TCTAAGTGTGTGAGGTATTTTCTCAGTTCAAGTAGAGAGATTATGAATGAA	864				
Db	126051	TCTAAGTGTGTGAGGTATTTTCTCAGTTCAAGTAGAGAGATTATGAATGAA	125992				
QY	8						

QY	925	GGGATTCCTCGGGGGCTCGTCTGTATATATCAAGAGTATGTCACAAATCTTACATAGAG	984
Db	125931	GGGATTCCTTCGGGGGGTCTCTGTATATATCAAGAGTATGTCACAAATCTTACATAGAG	125872
QY	985	GTCACTGAGGAGGACCCAGGAGCTACTGCTGCCACAGGAGTATATTTGTACAAAAGCAA	1044
Db	125871	GTCACTGAGGAGGACCCAGGAGCTACTGCTGCCACAGGAGTATATTTGTAGAAAAGCAA	125812
QY	1045	CTCCCTCAGTCCACGCTGTTTAGAGCTGACACACCATCTCTATTTGTATCAGGAGGAT	1104
Db	125811	CTCCCTCAGTCCACGCTGTTTAGAGCTGACACACCATCTCTATTTGTATCAGGAGGAT	125752
QY	1105	GACATCATCTTATTCAGTGGCAAAATTTCTTGCCCTTGA	1143
Db	125751	GACATCATCTTATTCAGTGGCAAAATTTCTTGCCCTTGA	125713

  

RESULT 8	AC027524	189092 bp	DNA	linear	HTG 27-MAR-2003
LOCUS	AC027524	189092 bp	DNA	linear	HTG 27-MAR-2003
DEFINITION	Human sapiens chromosome 18 clone RP11-317G1 map 18, 3 unordered pieces.				
ACCESSION	AC027524	4	GI:21307437		
VERSION	HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.				
KEYWORDS	Human sapiens (human)				
SOURCE	Human sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (Bases 1 to 189092)				
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.				
TITLE	Human sapiens chromosome 18, clone RP11-317G1				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 189092)				
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,P., Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,K., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., McDermid,J., Meneus,L., Mihova,T., Miranda,C., Mianga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Plerer,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA				
REFERENCE	3 (bases 1 to 189092)				
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mianga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,				

RESULT 10							PAT 10-APR-2003
LOCUS	AR270261					DNA	
DEFINITION	Sequence 824 from patent US 6500938.						linear
ACCESSION	AR270261						
VERSION	AR270261.1	GI:29701495					
KEYWORDS	.						
SOURCE	Unknown.						
ORGANISM	Unknown:						
REFERENCE	Unclassified.						
AUTHORS	1 (bases 1 to 316)						
TITLE	Au-Young,J. and Sellhammer,J.J.						
JOURNAL	Composition for the detection of signaling pathway gene expression						
FEATURES	Patent: US 6500938-A 824 31-DEC-2002;						
	Location/Qualifiers						
	1..316						
BASE COUNT	92 a	63 c	72 g	88 t			1 others
	/organism="unknown"						

ORGANISM					
Unknown.					
Unclassified.					
1 (bases 1 to 600)					
Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,					
Carter,D., Retter,M.W., Mannion,J., Fan,L. and wang,A.					
TITLE Compositions and methods for the therapy and diagnosis of lung cancer					
JOURNAL Patent: US 6509448-A 233 21-JAN-2003;					
FEATURES					
source					
BASE COUNT 198 a 106 c 135 g 161 t					
ORIGIN					
Query Match 10.6%; Score 121; DB 6; Length 600;					
Best Local Similarity 100.0%; Pred.No.1.8e-57; Indels 0; Gaps 0;					
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY 1023 AAGTAATATTGTAGAAGCAACTCCCTCAGTCACGCTGTTTAGAGCTGACCACCAT 1082					
DB 600 AAGTAAATTGTAGAAGCAACTCCCTCAGTCACGCTGTTTAGAGCTGACCACCAT 541					
QY 1083 CCTATTGGTTATCAGGAAGGATGCATCATCTTATTCAGTGCGCAAGTTCTTGCCCTTG 1142					
DB 540 CCTATTGGTTATCAGGAAGGATGCATCATCTTATTCAGTGCGCAAGTTCTTGCCCTTG 481					
QY 1143 A 1143					
DB 480 A 480					
RESULT 13					
AX062606/c LOCUS 600 bp DNA linear PAT 24-JAN-2001					
DEFINITION Sequence 233 from Patent WO0100828.					
ACCESSION AX062606					
VERSION AX062606.1 GI:12540481					
KEYWORDS Homo sapiens (human)					
SOURCE Homo sapiens					
ORGANISM Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE 1					
AUTHORS Wang T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,					
Carter, D., Retter, M.W. and Mannion, J.					
TITLE Compositions and methods for the therapy and diagnosis of lung cancer					
JOURNAL Patent: WO 0100828-A 233 04-JAN-2001;					
CORIXA CORPORATION [US]					
FEATURES					
source					
BASE COUNT 198 a 106 c 135 g 161 t					
ORIGIN					
Query Match 10.6%; Score 121; DB 6; Length 600;					
Best Local Similarity 100.0%; Pred.No.1.8e-57; Indels 0; Gaps 0;					
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY 1023 AAGTAAATATTGTAGAAGCAACTCCCTCAGTCACGCTGTTTAGAGCTGACCACCAT 1082					
DB 600 AAGTAAATATTGTAGAAGCAACTCCCTCAGTCACGCTGTTTAGAGCTGACCACCAT 541					
QY 1083 CCTATTGGTTATCAGGAAGGATGCATCATCTTATTCAGTGCGCAAGTTCTTGCCCTTG 1142					
DB 540 CCTATTGGTTATCAGGAAGGATGCATCATCTTATTCAGTGCGCAAGTTCTTGCCCTTG 481					
QY 1143 A 1143					
DB 480 A 480					

Query Match	20.1%; Score 230; DB 6; Length 316;
Best Local Similarity	99.6%; Pred. No. 2.9e-120;
Matches	280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	822 ATTTTTCCTCAGTTCAGAGTAGAGAAAGATTATGAAATGAAACATATTTGAGAGCCCT 881
DB	1 ATTTTTCCTCAGTTCAGAGTAGAGAAATATGAAATGAAACATATTTGAGAGCCCT 60
QY	882 AGCGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGG 941
DB	61 AGCGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGG 120
QY	942 TCGTCTGTATATATCAAGGATGATCCAAATCTTACATAGAGTCTCAGGAGGGCAC 1001
DB	121 TCGTCTGTATATATCAAGGATGATCCAAATCTTACATAGAGTCTCAGGAGGGCAC 180
QY	1002 CGAGGCTACTGCTCCCAAGGAATATTTGTAGAAAGCACTCCCTCAGTCCACGCT 1061
DB	181 CGAGGCTACTGCTCCCAAGGAATATTTGTAGAAAGCACTCCCTCAGTCCACGCT 240
QY	1062 GTTTAGAGTGCACACCATCTCTATTTGTTATCNGAAGG 1102
DB	241 GTTTAGAGTGCACACCATCTCTATTTGTTATCNGAAGG 281
RESULT 11	
LOCUS	AR272521/c 600 bp DNA linear PAT 10-APR-2003
DEFINITION	Sequence 233 from patent US 6504010.
ACCESSION	AR272521
VERSION	AR272521.1 GI:29704406
KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 600)
AUTHORS	Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedwick, T.S., Carter, D., Retter, M.W., Mannion, J. and Fan, L.
TITLE	Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL	Patent: US 6504010-A 233 07-JAN-2003;
FEATURES	Location/Qualifiers
source	1..600
BASE COUNT	198 a 106 c 135 g 161 t
ORIGIN	/organism="unknown"
Query Match	10.6%; Score 121; DB 6; Length 600;
Best Local Similarity	100.0%; Pred. No. 1.8e-57;
Matches	121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1023 AAGTAATATTGTAGAAAGCACTCCCTCAGTCCAGCGCTGTTAGAGCTGACCCACCT 1082
DB	600 AAGTAATATTGTAGAAAGCACTCCCTCAGTCCAGCGCTGTTAGAGCTGACCCACCT 541
QY	1083 CCTATTGTTATCAGGAAGGATGACATCATCTATTTCAGTGGCAAGTTCTTGCCCTTG 1142
DB	540 CCTATTGTTATCAGGAAGGATGACATCATCTATTTCAGTGGCAAGTTCTTGCCCTTG 481
QY	1143 A 1143
DB	480 A 480
RESULT 12	
LOCUS	AR276102/c 600 bp DNA linear PAT 10-APR-2003
DEFINITION	Sequence 233 from patent US 6509448.
ACCESSION	AR276102
VERSION	AR276102.1 GI:29709747
KEYWORDS	Unknown.
SOURCE	Unknown.



```

/mol_type="mRNA"
/db_xref="taxon:10116"
/cell_type="mesangial"
1. .1143
/notes="human megin homolog"
/codon_start=1
/product="megsin"
/protein_id="BAJ16769.1"
/db_xref="GI:16269521"
/translat="MASHAANAEFGDLEFMDSSQNGNVPFSLISFTALSILPL
GARGCARQIDKALHFISPRQGNSSNQLQYOLKRVLADINSHDKYIELSTANGV
FAEKVPDFHKSVMCAENLNAKVERVDFTNDIQETRFKINKWIENHKGKVLGD
SSLSGAVVAVLVNAVYFKGKGAFTKSDTLSCFRSPGPGKAVNAHQRRNLEST
IQEPPQILLOVHOGISWIMLPEDDLSEIRSLSFQNLMDWTNSRKMKSQYVNVFL
PQFKLEKDMASHLKSGLVEDIFVSRADLSGLASGRLYVSKLMHKSLEIYESEET
EATAATESNIVEKLLPSTVFRADRFLEFVIRKNGILFTGKVSCEP"
BASE COUNT      345 a      232 c      261 g      305 t
ORIGIN
Query Match      2.7%; Score 31; DB 10; Length 1143;
Best Local Similarity 100.0%; Pred.No.1.3e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      437 TTGAAATGAAACACATGGCAAAATCAAGAA 467
      |||||||
DB      437 TTGAAATGAAACACATGGCAAAATCAAGAA 467

```

Search completed: December 14, 2003, 06:01:42  
Job time : 4377 secs

```

CDS
/mol_type="mRNA"
/db_xref="taxon:10116"
/cell_type="mesangial"
1. .1143
/notes="human megin homolog"
/codon_start=1
/product="megsin"
/protein_id="BAJ16769.1"
/db_xref="GI:16269521"
/translat="MASHAANAEFGDLEFMDSSQNGNVPFSLISFTALSILPL
GARGCARQIDKALHFISPRQGNSSNQLQYOLKRVLADINSHDKYIELSTANGV
FAEKVPDFHKSVMCAENLNAKVERVDFTNDIQETRFKINKWIENHKGKVLGD
SSLSGAVVAVLVNAVYFKGKGAFTKSDTLSCFRSPGPGKAVNAHQRRNLEST
IQEPPQILLOVHOGISWIMLPEDDLSEIRSLSFQNLMDWTNSRKMKSQYVNVFL
PQFKLEKDMASHLKSGLVEDIFVSRADLSGLASGRLYVSKLMHKSLEIYESEET
EATAATESNIVEKLLPSTVFRADRFLEFVIRKNGILFTGKVSCEP"
BASE COUNT      345 a      232 c      261 g      305 t
ORIGIN
Query Match      2.7%; Score 31; DB 10; Length 1143;
Best Local Similarity 100.0%; Pred.No.1.3e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      437 TTGAAATGAAACACATGGCAAAATCAAGAA 467
      |||||||
DB      437 TTGAAATGAAACACATGGCAAAATCAAGAA 467

```

```

Query Match      10.6%; Score 121; DB 6; Length 600;
Best Local Similarity 100.0%; Pred.No.1.8e-57;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1023 AAGTAAATTTCTAGAAAGCAACTCCCTCAGTCACCGCTGTAGAGCTGACACCCATT 1082
      |||||||
DB      600 AAGTAAATTTCTAGAAAGCAACTCCCTCAGTCACCGCTGTAGAGCTGACACCCATT 541
QY      1083 CCTATTCTTTCAGGAGGATGACATCACTTATTCAGTCGCAAGTTCTTGCCCTTG 1142
      |||||||
DB      540 CCTATTCTTTCAGGAGGATGACATCACTTATTCAGTCGCAAGTTCTTGCCCTTG 481
QY      1143 A 1143
DB      480 A 480

```

```

RESULT 15
AF105329      1143 bp      mRNA      linear      ROD 19-OCT-2001
LOCUS      Rattus norvegicus megin mRNA, complete cds.
DEFINITION      Rattus norvegicus megin mRNA, complete cds.
ACCESSION      AF105329
VERSION      AF105329.1 GI:16269520
KEYWORDS
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE      1 (bases 1 to 1143)
AUTHORS      Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T.,
      Wada T., Yagi M., Nagano N., Inagi R. and Kurokawa K.
TITLE      Cloning of rodent megin revealed its up-regulation in
      mesangio proliferative nephritis
JOURNAL      Kidney Int. 60 (2), 641-652 (2001)
MEDLINE      21368006
PUBMED      11473647
REFERENCE      2 (bases 1 to 1143)
AUTHORS      Miyata T., Nangaku T., Inagi R. and Kurokawa K.
TITLE      Direct Submission
JOURNAL      Submitted (09-NOV-1998) Institute of Medical Sciences, Tokai
      University, Bohseidai, Isehara, Kanagawa 259-1193, Japan
FEATURES
SOURCE      1. .1143
      /organism="Rattus norvegicus"

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

ON nucleic - nucleic search, using sw model

Run on: December 14, 2003, 03:19:41 ; Search time 377 Seconds  
(without alignments)  
8184.238 Million cell updates/sec

Title: US-09-508-997A-1

Perfect score: 1143

Sequence: 1 atggctccctgtcgcgcg.....gcaagtttttgccttga 1143

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_15Jun03.\*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1143	100.0	1143	20	AAK56712 Human megasin codin
2	1143	100.0	1143	21	AAK56712 Human megasin nucle
3	1143	100.0	1143	22	AAK56712 Human megasin nucle
4	1143	100.0	1950	15	AAQ56670 Human megakaryocyt
5	1143	100.0	2249	25	ABX76260 Lung cancer-associ
6	254	22.2	254	25	ACA56060 Human signalling p
7	230	20.1	316	25	ACA56226 Human signalling p
8	121	10.6	600	22	AAF68315 Human lung tumour

C	9	121	10.6	600	24	ABK38226	cDNA encoding clon
C	10	121	10.6	600	25	ACA10555	Human lung cancer-
C	11	121	10.6	600	25	ABX99506	Lung cancer therap
C	12	31	2.7	42	20	AAK56727	WO9915652 primer 1
	13	31	2.7	1147	20	AAK56714	Mouse megasin DNA.
	14	31	2.7	1147	21	AAK56714	Mouse megasin nucle
	15	31	2.7	1147	22	AAK56714	Mouse megasin nucle
	16	31	2.7	1229	20	AAK56713	Rat megasin DNA. R
	17	31	2.7	1229	21	AAK56713	Rat megasin nucleot
	18	31	2.7	1229	22	AAH48181	Rat megasin protein
	19	31	2.7	1229	22	AAK56713	Rat megasin nucleot
	20	30	2.6	30	22	AAK56713	PCR primer hM8-1.
	21	30	2.6	44	20	AAK56713	WO9915652 primer 3
	22	30	2.6	44	22	AAK56713	Human megasin prime
	23	29	2.5	38	20	AAK56726	WO9915652 primer 1
C	24	26	2.3	26	22	AAK56726	Human megasin prime
	25	25	2.2	37	15	AAQ56668	Human megakaryocyt
	26	23	2.0	23	15	AAQ56651	Human megakaryocyt
	27	23	2.0	29	20	AAK56722	WO9915652 primer 8
C	28	22	1.9	22	20	AAK56750	WO9915652 primer 3
	29	22	1.9	29	20	AAK56752	Human megakaryocyt
C	30	21	1.8	21	15	AAQ56652	Human megakaryocyt
	31	21	1.8	21	15	AAQ56653	Human megakaryocyt
C	32	21	1.8	21	15	AAQ56654	Human megakaryocyt
	33	21	1.8	21	15	AAQ56655	Human megakaryocyt
C	34	21	1.8	21	15	AAQ56655	Human megakaryocyt
	35	21	1.8	21	20	AAK56749	WO9915652 primer 3
	36	21	1.8	21	20	AAK56724	WO9915652 primer 1
C	37	21	1.8	27	20	AAK56748	WO9915652 primer 3
	38	21	1.8	30	20	AAK56751	WO9915652 primer 3
C	39	21	1.8	30	20	AAK56751	WO9915652 primer 9
	40	21	1.8	34667	24	AAK4328	Human transporter
	41	21	1.8	249999	25	ABZ80239	Human transmembrin ge
	42	20	1.7	343	25	ABX20549	Human GDP-mannose
	43	20	1.7	1050	20	AAK61596	B. burgdorferi ant
	44	20	1.7	1152	20	AAK61595	B. burgdorferi ant
C	45	20	1.7	4493	23	ABL28683	Drosophila melanog

#### ALIGNMENTS

RESULT 1  
AAK56712  
ID AAK56712 standard; DNA; 1143 BP.

XX AAK56712;  
XX AC AAK56712;  
XX DT 14-JUN-1999 (first entry)  
XX DE Human megasin coding region.  
XX DE Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;  
KW human; rat; murine; ss.  
XX CS Homo sapiens.  
XX PN WO9915652-AL.  
XX PD 01-APR-1999.  
XX PF 22-SEP-1998; 98WO-JP04269.  
XX PR 22-SEP-1997; 97JP-0275302.  
XX PA (KURO/) KUROKAWA K.  
XX PA (MIYA/) MIYATA T.  
XX PI Miyata T;  
XX DR WPI; 1999-275983/23.  
XX DR P-PSDB; AAY08254.

PT Megin protein expressed specifically in mesangial cells

PS Claim 4; Page 58-61; 100pp; Japanese.

XX This invention describes the isolation of novel megin nucleic acid and  
CC proteins from human, rat and mouse tissue. This protein is expressed  
CC specifically in mesangial cells. The products of the invention are  
CC useful for the treatment and diagnosis of diseases involving mesangial  
CC cells, such as IgA nephropathy.

XX SQ Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;

Query Match 100.0%; Score 1143; DB 20; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCCTCCCTGCTGAGCAATGAGAGTTTGGCTCAACCTGTCAGAGAGATGGAT 60  
Db 1 ATGCCCTCCCTGCTGAGCAATGAGAGTTTGGCTCAACCTGTCAGAGAGATGGAT 60  
Qy 61 GACAAATCAAGAAATGGAATGTTCTTTCTCTCTGAGCCCTCTGCTGCCCTGGCC 120  
Db 61 GACAAATCAAGAAATGGAATGTTCTTTCTCTCTGAGCCCTCTGCTGCCCTGGCC 120  
Qy 121 CTGCTCCGCTGGGCGCTCAAGATGACTCCCTCTCTGAGATGATAGTTGCTTCATGTT 180  
Db 121 CTGCTCCGCTGGGCGCTCAAGATGACTCCCTCTCTGAGATGATAGTTGCTTCATGTT 180  
Qy 181 AACACTGCTCAGATATGGAATCTCTCTATAGTCACTGAGGCTCCAGTCTCAACTG 240  
Db 181 AACACTGCTCAGATATGGAATCTCTCTATAGTCACTGAGGCTCCAGTCTCAACTG 240  
Qy 241 AAAGAGTTTTTCTGATATAATGCAATCCCAAGGATATGATCTCAGCATTTGCAAT 300  
Db 241 AAAGAGTTTTTCTGATATAATGCAATCCCAAGGATATGATCTCAGCATTTGCAAT 300  
Qy 301 GGGCTTTTCTGAAAAGTGTATGGCTTTTCAAGGCTACATGAGTGTCCGAAAAA 360  
Db 301 GGGCTTTTCTGAAAAGTGTATGGCTTTTCAAGGCTACATGAGTGTCCGAAAAA 360  
Qy 361 TTATACAGTCCAAAGTGGAGGAGTTCAGTTTACGATCAATTTAGAGACACTAGAGT 420  
Db 361 TTATACAGTCCAAAGTGGAGGAGTTCAGTTTACGATCAATTTAGAGACACTAGAGT 420  
Qy 421 AATATTAATAGTGGTGTGAAAATGAAACACATGCGCAAAATCAAGAACCTGATCGTAA 480  
Db 421 AATATTAATAGTGGTGTGAAAATGAAACACATGCGCAAAATCAAGAACCTGATCGTAA 480  
Qy 481 GGTGCAATAGCTCATCTCTCTAATGGTGTGGTATGCTGTGCTACCTCAAGGCAAG 540  
Db 481 GGTGCAATAGCTCATCTCTCTAATGGTGTGGTATGCTGTGCTACCTCAAGGCAAG 540  
Qy 541 TGGCAATCAGCTTCCCAAGAGCGAATCAATTAATGCAATTTCAATCTCCCAAGTGC 600  
Db 541 TGGCAATCAGCTTCCCAAGAGCGAATCAATTAATGCAATTTCAATCTCCCAAGTGC 600  
Qy 601 TCTGGGAAGGAGTGGCCATGATGATCAGGAACGGAAGTTCAATTTGCTGTATTTAG 660  
Db 601 TCTGGGAAGGAGTGGCCATGATGATCAGGAACGGAAGTTCAATTTGCTGTATTTAG 660  
Qy 661 GACCAATCAATGAGATTTCTGAGTCAATCAATGAGTGGGATTAACATGATGCTCTG 720  
Db 661 GACCAATCAATGAGATTTCTGAGTCAATCAATGAGTGGGATTAACATGATGCTCTG 720  
Qy 721 CTGCTGAGATGACCTCTGGAATGAAACAAATGACCTTTCAAGATCTAATGGAA 780  
Db 721 CTGCTGAGATGACCTCTGGAATGAAACAAATGACCTTTCAAGATCTAATGGAA 780  
Qy 781 TGGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTATTTTTTCTCAGTTCAAG 840  
Db 781 TGGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTATTTTTTCTCAGTTCAAG 840  
Qy 841 ATAGAGAGAAATATGAATGAACAAATATTTGAGCCCTAGGCTGAAAGATATCTTT 900

RESULT 2

AAA9294  
ID AAA9294 standard; cDNA; 1143 BP.

XX AAA9294;

XX 30-JAN-2001 (first entry)

XX Human megin nucleotide sequence SEQ ID NO:1.

XX Megin; mesangium-predominant gene; serpin regulated; nephropathy;  
KW IGA; immunoglobulin A; detection; renal function; renal disorder;  
KW diagnosis; biological sample; blood; urine; ss.

XX Homo sapiens.

XX WO200057189-A1.

XX 28-SEP-2000.

XX 17-MAR-2000; 2000WO-JP01646.

XX 19-MAR-1999; 99JP-0075305.

XX 28-OCT-1999; 99JP-0306623.

XX (KURO/) KUROKAWA K.

XX (FUSO) FUSO PHARM IND LTD.

XX (MIYA) MIYATA T.

XX Miyata T;

XX WPI; 2000-611642/58.

XX P-PSDB; AAB24142.

XX Evaluating renal function comprises assaying megin protein in  
XX biological sample

XX Example 5; Page 62-66; 93pp; Japanese.

XX The present invention describes a method for evaluating renal function.  
CC The method comprises assaying megin protein in biological sample. Also  
CC described are: (1) use of an anti-megin protein antibody for diagnosing  
CC renal function; and (2) a kit for detecting megin protein comprising:  
CC (a) anti-megin protein antibody attached to solid magnetic particles;  
CC (b) direct or indirect fixing for the antibody to the particles; and  
CC (c) a magnet. The process is useful for evaluating renal function and  
CC diagnosing renal disorders by assaying megin protein in biological  
CC samples (preferably urine or blood). The process is reproducible and  
CC gives accurate results. The present sequence encodes the human megin

CC protein, which is given in the exemplification of the present invention.  
XX  
SQ Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;  
Query Match 100.0%; Score 1143; DB 21; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGSCCTCCCTTCTGTCGACAAATGACAGATTTTCTTCAACCTGTTACAGAGATGAT 60  
DB 1 ATGSCCTCCCTTCTGTCGACAAATGACAGATTTTCTTCAACCTGTTACAGAGATGAT 60  
QY 61 GACATCAAGAAATGGAATGCTTCTTCTCTCTGAGCCTCTTCTGCTCCCTGGCC 120  
DB 61 GACATCAAGAAATGGAATGCTTCTTCTCTCTGAGCCTCTTCTGCTCCCTGGCC 120  
QY 121 CTGCTCCGCTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGCTTCATGTT 180  
DB 121 CTGCTCCGCTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGCTTCATGTT 180  
QY 181 AACACTGCTCAGATATGAAATCTTCTTAATAGTCACTCAGGCTCCAGTCTCAACTG 240  
DB 181 AACACTGCTCAGATATGAAATCTTCTTAATAGTCACTCAGGCTCCAGTCTCAACTG 240  
QY 241 AAAAGATTTTCTGATATAATGCAATCCACAGAGATTAATGATCTCAGATGCTGAAT 300  
DB 241 AAAAGATTTTCTGATATAATGCAATCCACAGAGATTAATGATCTCAGATGCTGAAT 300  
QY 301 GGGCTTTTCTGAAAGATGATGGCTTTTCTAAGAGACTACATTCAGTGTGCCAAAAA 360  
DB 301 GGGCTTTTCTGAAAGATGATGGCTTTTCTAAGAGACTACATTCAGTGTGCCAAAAA 360  
QY 361 TTATACGATGCAAAATGAGGAGTTGACTTTTACAGATCAATTTAGAGACTACAGT 420  
DB 361 TTATACGATGCAAAATGAGGAGTTGACTTTTACAGATCAATTTAGAGACTACAGT 420  
QY 421 AATATATAGTGGTTCGAAATCAACACATGCGCAAAATCAAGACGTTGATGCTGAA 480  
DB 421 AATATATAGTGGTTCGAAATCAACACATGCGCAAAATCAAGACGTTGATGCTGAA 480  
QY 481 GGTGCAATAGCTCATCTGCTGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB 481 GGTGCAATAGCTCATCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 541 TGGCAATCAGCTTCCACAGAGCGAAACATTAATGCGATTTCAAAATCTCCAGATGC 600  
DB 541 TGGCAATCAGCTTCCACAGAGCGAAACATTAATGCGATTTCAAAATCTCCAGATGC 600  
QY 601 TCTGGGAAGGAGTCGCGCATGATGATCAGAGACGGAAGTTCATTTGCTGTTATTGAG 660  
DB 601 TCTGGGAAGGAGTCGCGCATGATGATCAGAGACGGAAGTTCATTTGCTGTTATTGAG 660  
QY 661 GACCCATCAATGAAGATTTCTGAGCTCAGATACAAATGTTGGCATAAATGTAAGTTCTG 720  
DB 661 GACCCATCAATGAAGATTTCTGAGCTCAGATACAAATGTTGGCATAAATGTAAGTTCTG 720  
QY 721 CTGCTGAGATGACCTCTGATATGAAACAACTGACCTTTCAGATCTAATGAA 780  
DB 721 CTGCTGAGATGACCTCTGATATGAAACAACTGACCTTTCAGATCTAATGAA 780  
QY 781 TGGCAATCAAGCGCAATGACCTTCTAAGTATGTTGAGTATTTTCTCAGTTCAAG 840  
DB 781 TGGCAATCAAGCGCAATGACCTTCTAAGTATGTTGAGTATTTTCTCAGTTCAAG 840  
QY 841 ATAGAGAAGATTTATGAATGAACAAATATTTAGAGCCCTAGGCTGAAAGATATCTTT 900  
DB 841 ATAGAGAAGATTTATGAATGAACAAATATTTAGAGCCCTAGGCTGAAAGATATCTTT 900  
QY 901 GATGAATCAAGAGATCTCTGAGATGCTCTGAGGCTGCTGAGGCTGCTGATATCAAG 960  
DB 901 GATGAATCAAGAGATCTCTGAGATGCTCTGAGGCTGCTGAGGCTGCTGATATCAAG 960  
QY 961 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGCGACCGAGGCTACTGCTGCCACA 1020

DB 961 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGCGACCGAGGCTACTGCTGCCACA 1020  
QY 1021 GGAAGTAAATTTAGAAAAGCAACTCCCTCAGTCCAGCTGCTGTTAGAGCTGACACCCA 1080  
DB 1021 GGAAGTAAATTTAGAAAAGCAACTCCCTCAGTCCAGCTGCTGTTAGAGCTGACACCCA 1080  
QY 1081 TTCTTATTTCTATCAGGAGGAGTACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCCT 1140  
DB 1081 TTCTTATTTCTATCAGGAGGAGTACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCCT 1140  
QY 1141 TGA 1143  
DB 1141 TGA 1143  
RESULT 3  
AAF82438  
ID AAF82438 standard; cDNA; 1143 BP.  
XX  
AC AAF82438;  
XX  
DT 10-JUL-2001 (first entry)  
XX  
DE Human meglin nucleotide sequence.  
XX  
KW Human; meglin; mesangial cell proliferative nephritis; nephrotropic;  
KW transgenic mouse; glomerular disease; animal model; drug screening; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1143  
FT /tag= a  
FT /product= "meglin"  
XX  
PN WO200124628-A1.  
XX  
PD 12-APR-2001.  
XX  
PF 06-OCT-2000; 2000WO-JP06988.  
XX  
PR 06-OCT-1999; 99JP-0285736.  
XX  
PA (KURO/) KUROKAWA K.  
PA (MIYA/) MIYATA T.  
PI Miyata T;  
XX  
XX WPI: 2001-300136/31.  
DR P-PSDB; AAB83075.  
XX  
PT Mouse model for mesangial cell proliferative nephritis for development  
PT and screening of new treatments -  
XX  
PS Claim 2; Page 42-44; 62pp; Japanese.  
XX  
CC The present sequence encodes human meglin. The human meglin coding  
CC sequence may be introduced into a mouse to produce an animal model of  
CC mesangial cell proliferative nephritis. The symptoms include  
CC enlargement of the mesangial base region, sedimentation of an immune  
CC complex and an increase in mesangial cells. The animal model is useful  
CC for analysing the pathology of chronic glomerular diseases and for  
CC screening compositions for prevention and treatment of the diseases.  
CC Highly uniform models can be made easily and in large numbers using  
CC this method.  
XX  
SQ Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;  
Query Match 100.0%; Score 1143; DB 22; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTCCCTTGTGTCAGCAAAATGAGATTTGGCTTCAACCTGTTTCAGAGATGGAT 60  
Db 1 ATGGCTCCCTTGTGTCAGCAAAATGAGATTTGGCTTCAACCTGTTTCAGAGATGGAT 60  
QY 61 GACAAATCAAGAAATGAAATGTGTTCTTTTCCCTCTCTGAGCCCTTGGCTGCCGCC 120  
Db 61 GACAAATCAAGAAATGAAATGTGTTCTTTTCCCTCTCTGAGCCCTTGGCTGCCGCC 120  
QY 121 CTGGTCGGCTTGGGGCTCAAGATGACTCCCTCTCTCAGATTGATAAGTTGCTTCATGTT 180  
Db 121 CTGGTCGGCTTGGGGCTCAAGATGACTCCCTCTCTCAGATTGATAAGTTGCTTCATGTT 180  
QY 181 AACACTGCTCAGGATATGAAATCTTCTTAATAGTCAGTCAGGGCTCCAGTCTCAACTG 240  
Db 181 AACACTGCTCAGGATATGAAATCTTCTTAATAGTCAGTCAGGGCTCCAGTCTCAACTG 240  
QY 241 AAAGAGTTTTTCTGATATATAATATGATCCCAAGATATGATCTCAGCATGTGAAT 300  
Db 241 AAAGAGTTTTTCTGATATATAATATGATCCCAAGATATGATCTCAGCATGTGAAT 300  
QY 301 GGGCTTTTGTGTAAGATGATGCTTTGATTAAGCACTACATTCAGTGTGCCGAAAA 360  
Db 301 GGGCTTTTGTGTAAGATGATGCTTTGATTAAGCACTACATTCAGTGTGCCGAAAA 360  
QY 361 TTATACGATGCCAAAGTGGAGCGAGTTGACTTTAGCAATCATTTAGAAAGCACTAGACT 420  
Db 361 TTATACGATGCCAAAGTGGAGCGAGTTGACTTTAGCAATCATTTAGAAAGCACTAGACT 420  
QY 421 AATATTAATAGTGGTGAATATGAATGAACACATGCCAAATCAAGACGTGATGGTGA 480  
Db 421 AATATTAATAGTGGTGAATATGAATGAACACATGCCAAATCAAGACGTGATGGTGA 480  
QY 481 GGTGGATAGCTCACTGCTGTAATGGTGGTGGTGAATGCTGCTACTTCAAGGCAAG 540  
Db 481 GGTGGATAGCTCACTGCTGTAATGGTGGTGGTGAATGCTGCTACTTCAAGGCAAG 540  
QY 541 TGGCAATCAGCTTCCACCAAGCGAAACATTAATGCGAATTCCTCAATCTCCCAAGTGC 600  
Db 541 TGGCAATCAGCTTCCACCAAGCGAAACATTAATGCGAATTCCTCAATCTCCCAAGTGC 600  
QY 601 TCTGGAGAGCTGCGCATGATGCTCAGTCAAGACCGAGTTCATTTGCTGTTATTGAG 660  
Db 601 TCTGGAGAGCTGCGCATGATGCTCAGTCAAGACCGAGTTCATTTGCTGTTATTGAG 660  
QY 661 GACCATCAATGAAGATCTTCTGAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 661 GACCATCAATGAAGATCTTCTGAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
QY 721 CTGCTGAGATGACCTCTCTGAAATGAAACAACTGACCTTTCAGATCTCAAG 840  
Db 721 CTGCTGAGATGACCTCTCTGAAATGAAACAACTGACCTTTCAGATCTCAAG 840  
QY 781 TGGACCAATCAGGCGAATGACCTCTGAAATGAAACAACTGACCTTTCAGATCTCAAG 840  
Db 781 TGGACCAATCAGGCGAATGACCTCTGAAATGAAACAACTGACCTTTCAGATCTCAAG 840  
QY 841 ATAGAGAGATTAATGAATGAACAACTGAGGCTTTCAGATCTCAAGATCTT 900  
Db 841 ATAGAGAGATTAATGAATGAACAACTGAGGCTTTCAGATCTCAAGATCTT 900  
QY 901 GATGAATCAAGAGATCTCTCTGGGATTCCTTGGGGGCTGCTGCTGATATCAAGG 960  
Db 901 GATGAATCAAGAGATCTCTCTGGGATTCCTTGGGGGCTGCTGCTGATATCAAGG 960  
QY 961 ATGATCCCAATCTTACATGAGGCTGCTGAGGAGGCGACCGAGGCTTCTGCTGCA 1020  
Db 961 ATGATCCCAATCTTACATGAGGCTGCTGAGGAGGCGACCGAGGCTTCTGCTGCA 1020  
QY 1021 GGAAGTAATATTGAGAAAGCAATCTCCTCAGTCCACGCTGTTTTCAGAGTCAACCCA 1080  
Db 1021 GGAAGTAATATTGAGAAAGCAATCTCCTCAGTCCACGCTGTTTTCAGAGTCAACCCA 1080  
QY 1081 TTCTTATTTATCAGGAGGATGACATCTTATTCAGTGGCAAGTTTCTTGGCTT 1140

1081 TTCTTATTTATCAGGAGGATGACATCTTATTCAGTGGCAAGTTTCTTGGCTT 1140  
QY 1141 TGA 1143  
Db 1141 TGA 1143

RESULT 4  
AAQ56670  
ID AAQ56670 standard; cDNA; 1950 BP.  
XX  
AC AAQ56670;  
XX  
XX 25-MAR-2003 (updated)  
DT 16-AUG-1994 (first entry)  
XX  
XX Human megakaryocyte differentiation factor coding sequence.  
XX  
XX Human megakaryocyte differentiation factor; MDP; thrombopoietin;  
KW haematopoietic stimulating factor; thrombocytopoiesis; platelet;  
XX bone marrow transplantation; cancer chemotherapy; ds.  
XX  
OS Homo sapiens.  
XX  
XX Location/Qualifiers  
FH Key 74..1216  
FT CDS /\*tag= a  
PT /product= megakaryocyte\_differentiation\_factor  
XX  
XX EP583884-A1.  
XX  
XX 23-FEB-1994.  
XX  
XX 19-JUL-1993; 93EP-0305654.  
XX  
XX 17-JUL-1992; 92JP-0212305.  
PR 04-MAR-1993; 93JP-0067339.  
XX  
XX (SUNR ) SUNTORY LTD.  
PA (TSUJ/) TSUJIMOTO M.  
XX  
XX Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;  
PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamaichi K;  
XX  
XX WPI; 1994-058782/08.  
DR P-PSDB; AAR48379.  
XX  
XX New megakaryocyte differentiation factor - isolated from human  
PT epidermoid carcinoma cells, used to treat conditions involving a  
PT decrease in platelets  
XX  
XX Claim 13; Page 30-32; 47pp; English.  
XX  
XX The cDNA coding for human MDP was isolated from a library prepared  
CC using mRNA derived from human epidermoid carcinoma A431 cells. Human  
CC MDP stimulates differentiation of megakaryocytes from myeloid cells  
CC in the presence of Ii-3. The MDP acts in vivo as a thrombopoietin  
CC making it useful for treatment of diseases involving a decrease in  
CC platelet number (esp. thrombocytopaenia) such as occurs in bone  
CC marrow transplantation and in chemotherapy.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 1950 BP; 570 A; 384 C; 407 G; 589 T; 0 other;  
SQ

Query Match 100.0%; Score 1143; DB 15; Length 1950;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTCCCTTGTGTCAGCAAAATGAGATTTGGCTTCAACCTGTTTCAGAGATGGAT 60  
Db 74 ATGGCTCCCTTGTGTCAGCAAAATGAGATTTGGCTTCAACCTGTTTCAGAGATGGAT 133

Db 1214 TGA 1216

RESULT 5

ABX76260

ID ABX76260 standard; DNA; 2249 BP.

XX

AC ABX76260;

DT 02-APR-2003 (first entry)

XX

DE Lung cancer-associated polynucleotide #126.

XX

KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema; antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis; small cell lung cancer; benign lesion; precancerous lesion; bronchitis; chronic obstructive pulmonary disease; hypersensitivity pneumonitis; interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX

XX Unidentified.

XX

XX WC200256443-A2.

PN

XX

PD 31-OCT-2002.

XX

XX

PF 18-APR-2002; 2002WO-US12476.

XX

PR 18-APR-2001; 2001US-284770P.

PR

PR 10-MAY-2001; 2001US-290492P.

PR

PR 09-NOV-2001; 2001US-339245P.

PR

PR 13-NOV-2001; 2001US-350666P.

PR

PR 29-NOV-2001; 2001US-334370P.

PR

XX 12-APR-2002; 2002US-372246P.

XX

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX

XX

PI Aziz N, Murray R;

XX

XX WPI; 2003-093161/08.

DR

DR P-PSDB; ABU56533.

XX

XX

PT Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer -

XX

PS Claim 22; Page 288; 453pp; English.

XX

XX The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the invention.

XX

XX

XX Sequence 2249 BP; 649 A; 452 C; 461 G; 683 T; 4 other;

XX

XX Query Match 100.0%; Score 1143; DB 25; Length 2249;

XX Best Local Similarity 100.0%; Fred. No. 0;

		Matches 1143;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	ATGGCTCCCTTGTGTCAGCAAAATCAGAGTTTGTCTTCAACCTGTTTCAGAGATGGAT	60							
Db	365	ATGGCTCCCTTGTGTCAGCAAAATCAGAGTTTGTCTTCAACCTGTTTCAGAGATGGAT	424							
QY	61	GACATCAAGGAATGAAATGTGTTTCTCTCTCTGAGCTTCTGCTGCCCTGCC	120							
Db	425	GACATCAAGGAATGAAATGTGTTTCTCTCTCTGAGCTTCTGCTGCCCTGCC	484							
QY	121	CTGGTCCGCTTGGGCGCTCAAGATCTCCTCTCTCAGATTGATGATTTCTCATGTT	180							
Db	485	CTGGTCCGCTTGGGCGCTCAAGATCTCCTCTCTCAGATTGATGATTTCTCATGTT	544							
QY	181	AACATCTCTCAGATGATGAAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCACTG	240							
Db	545	AACATCTCTCAGATGATGAAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCACTG	604							
QY	241	AAAGAGTTTCTTCTGATATAAATGATCCACCAAGGATTTATGATCTCAGATTTGTAAT	300							
Db	605	AAAGAGTTTCTTCTGATATAAATGATCCACCAAGGATTTATGATCTCAGATTTGTAAT	664							
QY	301	GGGCTTTTCTGAAAAAGTGTATGCTTTCTAAGACTACATTTAGCTGTCGCAAAA	360							
Db	665	GGGCTTTTCTGAAAAAGTGTATGCTTTCTAAGACTACATTTAGCTGTCGCAAAA	724							
QY	361	TTATAGATGCCAAGTGGAGCGAGTTGATCTTACGAATCANTTAGAGACACTAGACCT	420							
Db	725	TTATAGATGCCAAGTGGAGCGAGTTGATCTTACGAATCANTTAGAGACACTAGACCT	784							
QY	421	AATATTAATGAAGTGGTGTGAATGAAACACATGCAAAATCAAGAACGTGATTTGTA	480							
Db	785	AATATTAATGAAGTGGTGTGAATGAAACACATGCAAAATCAAGAACGTGATTTGTA	844							
QY	481	GGGCGATAGATCTATCTGCTGTAAGTGGTCTGCTGATGCTGTACTTCAAGGCAAG	540							
Db	845	GGGCGATAGATCTATCTGCTGTAAGTGGTCTGCTGATGCTGTACTTCAAGGCAAG	904							
QY	541	TGGCAATCAGCTTCCACCAAGGCAAAACCAATAATTCGCAATTTCAAAATCTCCCAAGTC	600							
Db	905	TGGCAATCAGCTTCCACCAAGGCAAAACCAATAATTCGCAATTTCAAAATCTCCCAAGTC	964							
QY	601	CTGCGAAGGCACTCCGCAATGATCANTCAGGAACGGAAGTTCAATTTGCTTTATTTAG	660							
Db	965	CTGCGAAGGCACTCCGCAATGATCANTCAGGAACGGAAGTTCAATTTGCTTTATTTAG	1024							
QY	661	GAACCATCAATGAAGATCTTTCAGCTCAGATCAATGCTGGGCAATAACATGATCTCTG	720							
Db	1025	GAACCATCAATGAAGATCTTTCAGCTCAGATCAATGCTGGGCAATAACATGATCTCTG	1084							
QY	721	CTGCTGAGATGACCTCTCTGAAATTTGAAAAAACAACCTGACCTTTCAGAAATCTAATGAA	780							
Db	1085	CTGCTGAGATGACCTCTCTGAAATTTGAAAAAACAACCTGACCTTTCAGAAATCTAATGAA	1144							
QY	781	TGGAACAATCCAGGCAATGACCTCTAAGTATGTTGAGTATTTTCTCAGTCTCAAG	840							
Db	1145	TGGAACAATCCAGGCAATGACCTCTAAGTATGTTGAGTATTTTCTCAGTCTCAAG	1204							
QY	841	ATAGAGAGATTTATGAATGAACCAATTTTGAAGAGCCCTAGGCTGAAAGATATCTTT	900							
Db	1205	ATAGAGAGATTTATGAATGAACCAATTTTGAAGAGCCCTAGGCTGAAAGATATCTTT	1264							
QY	901	GATCAATCCAAAGCAGATCTCTCTGGGATTCCTTTCGGGGGTCGCTGTATATATCAAGG	960							
Db	1265	GATCAATCCAAAGCAGATCTCTCTGGGATTCCTTTCGGGGGTCGCTGTATATATCAAGG	1324							
QY	961	ATGATCCAAATCTTATAGAGTCTATGAGAGGCAACGAGGCTACTGTGTCACA	1020							
Db	1325	ATGATCCAAATCTTATAGAGTCTATGAGAGGCAACGAGGCTACTGTGTCACA	1384							
QY	1021	GGAGTAAATTTGTAGAAAGCAATCTCCTCAGTCCAGCTGTTTAGAGCTGACCAACCA	1080							
Db	1385	GGAGTAAATTTGTAGAAAGCAATCTCCTCAGTCCAGCTGTTTAGAGCTGACCAACCA	1444							

QY	1081	TTCTATTTTGTATCAGAAAGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCT	1140
Db	1445	TTCTATTTTGTATCAGAAAGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCT	1504
QY	1141	TGA	1143
Db	1505	TGA	1507

RESULT 6

ACAS6060	ACA56060 standard; cDNA; 254 BP.
XX	ACA56060;
XX	06-JUN-2003 (first entry)
XX	Human signalling pathway polynucleotide probe SEQ ID NO 658.
XX	Human; probe; ss; array element; Parkinson's disease;
XX	signalling pathway population; cancer; adenocarcinoma; leukaemia;
XX	immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX	Homo sapiens.
XX	US6500938-B1.
XX	31-DEC-2002.
XX	30-JAN-1998; 98US-0016434.
XX	30-JAN-1998; 98US-0016434.
XX	(INCY-) INCYTE GENOMICS INC.
XX	Au-Young J, Seilhamer JU;
XX	WPI; 2003-352189/33.
XX	Combination of polynucleotide probes, useful as array elements in a
XX	microarray for monitoring the expression of a number of target
XX	polynucleotides
XX	Claim 1; SEQ ID NO 658; 65pp; English.

The invention relates to a combination which, comprises a number of polynucleotide probes comprising a sequence selected from one of the 1490 sequences mentioned in the specification. The combination is useful as an array element in a microarray for monitoring the expression of a number of target polynucleotides. The microarray is particularly useful in the diagnosis and treatment of cancer and immunopathology and neuropathology. The microarray is useful in diagnostics and treatment regimens, drug discovery and development, toxicological and carcinogenicity studies, forensics and pharmacogenomics. The microarray is also useful for monitoring progression of diseases and for developing sophisticated profiles for the effects of currently available therapeutic drugs. The combination is also useful for purifying a subpopulation of mRNAs, cDNAs and genomic fragments and in research and diagnostic applications. The array can detect changes in expression in a large number of genes coding for different signalling pathway populations which can be used to diagnose various diseases including cancer e.g. adenocarcinoma and leukaemia, immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease and Parkinson's disease. The present sequence represents a polynucleotide probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html?DocID=06500938B1](http://seqdata.uspto.gov/sequence.html?DocID=06500938B1).

Query Match	22.2%;	Score	254;	DB	25;	Length	254;
Best Local Similarity	100.0%;	Pred. No.	7.7e-123;				
Sequence	254 BP;	83 A;	38 C;	64 G;	69 T;	0 other;	

Query Match 20.1%; Score 230; DB 25; Length 116;  
Best Local Similarity 99.8%; Pred. No. 3.2e-110;  
Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
822 ATTCTTTCTCAGTTCAAGTATGAGAGAGATATCGAATGAACACATATTGTGAGCCCT 881

Db 1 A T T T T T T C C T C A G T C C A G A T A G A G A G A G A G A T T A T G A A T G A A C A A T A T T T G A G A G C C C T 60

Qy 882 A G G C T G A A A G A T A T C T T T G A T G A T C C A A G A G A T C T C T C T G G A T G T C T T C G G G G G 941

Ob 61 A G G G T T G A A A G A T A T C T T T G A T G A T C C A A G A G A T C T C T T G G G T G T C T T C G G G G G 120

Qy	942	TGTCGTATATATCAAGGATGACACAAATCTTACATAGAGCTCACTGAGGAGGGCAC	1001
Dy	121	TGTCGTATATATCAAGGATGACACAAATCTTACATAGAGGCTCACTGAGGAGGGCAC	180
Qy	1002	CGNGCTACTGTGCCACAGGAAGTATATCTAGAAAACCAATCCCTCAGTCCACGCT	1061
Dy	181	CGNGCTACTGTGCCACAGGAAGTATATCTAGAAAACCAATCCCTCAGTCCACGCT	240

1062 GTTTAGACTCACCACCATTCCTATTGTTATCAGNAG 1102  
 241 GTTTAGACTCACCACCATTCCTATTGTTATCAGNAG 281

RESULT 8  
 AAT68315/C  
 ID AAT68315 standard: cDNA: 600 bp.

XX	AMF68315;	
XX		
XX		
DT	12-APR-2001	(first entry)

XX Human lung tumour protein related nucleotide sequence SEQ ID NO:233 .  
DE  
XX  
XX Human; lung cancer; lung tumour; lung tumour protein; gene therapy;  
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;  
KW cytostatic; antisense inhibition; ss.

XX	XX	Homo sapiens.
OS	OS	
XX	XX	
PN	PN	WC20010828-A2.

XX		
PD	04-JAN-2001.	
XX		
EF	30-JUN-2000; 2000WC-US18061.	
XX		
PR	30-JUN-1999; 99US-0346492.	
PR	15-OCT-1999; 99US-0419356.	
PR	17-DEC-1999; 99US-0466867.	
PR	30-DEC-1999; 99US-0476300.	
PR	06-MAR-2000; 2000US-0519642.	
PR	22-MAR-2000; 2000US-0533077.	
PR	10-APR-2000; 2000US-0546259.	
PR	27-APR-2000; 2000US-0560406.	
PR	05-JUN-2000; 2000US-0589184.	
XX		
PA	(CORI-) CORIXA CORP.	
XX		
PI	Wang T, Bangur CS, Lodes MJ,	Fanger GR, Vedvick TS, Carter D;
PI	Retter MW, Mannion J;	
XX		

XX Wang T, Bangor CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;  
PI Retter MW, Mannion J;  
XX



DR WPI; 2001-071498/08.  
 XX Lung tumor-associated proteins and the nucleic acids that encode them,  
 PT useful for preventing, diagnosing and treating lung cancer -  
 XX  
 XX Example 1; Page 223; 436pp; English.  
 XX The present invention describes immunogenic portions of lung tumour-  
 CC associated proteins (I) and the nucleic acids (NAs) that encode them.  
 CC (I) have cytostatic activity and can be used in gene therapy, antisense  
 CC inhibition and in vaccines. The NAs and the lung tumour-associated  
 CC proteins they encode may be used in the prevention, treatment and  
 CC diagnosis of diseases associated with their inappropriate expression,  
 CC especially lung cancers. For example, the NAs may be administered to  
 CC treat diseases by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of the protein by expressing inactive proteins  
 CC or to supplement the patient's own production of (I). Additionally, the  
 CC NAs may be used to produce the lung-tumour associated protein, according  
 CC to standard recombinant DNA methodology. Conversely, antisense NA  
 CC molecules may be administered to down regulate protein expression by  
 CC binding with the cells own genes and preventing their expression. The NA  
 CC and complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar NA sequences in  
 CC samples, and hence which patients may be in need of treatment for lung  
 CC cancer. The (I) may be used as antigens in the production of antibodies  
 CC and in assays to identify modulators (agonists and antagonists) of the  
 CC expression and activity of the protein. AAF68083 to AAF68878 and  
 CC AAB76848 to AAB76878 represent human lung tumour protein related  
 CC nucleotide and protein sequences which are used in the exemplification  
 CC of the present invention.  
 XX  
 XX Sequence 600 BP; 198 A; 106 C; 135 G; 161 T; 0 other;  
 SQ  
 Query Match 10.6%; Score 121; DB 22; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-53;  
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1023 AAGTAAATTTGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCCATT 1082  
 Db 600 AAGTAAATTTGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCCATT 541  
 QY 1083 CCTATTGTTATCAGGAAGATGACATCATCTTATTTCAGTGGCAAGTTCTTGCCCTTG 1142  
 Db 540 CCTATTGTTATCAGGAAGATGACATCATCTTATTTCAGTGGCAAGTTCTTGCCCTTG 481  
 QY 1143 A 1143  
 Db 480 A 480  
 RESULT 9  
 ABK38226/c  
 ID ABK38226 standard; cDNA; 600 BP.  
 XX  
 XX AC ABK38226;  
 XX  
 XX 21-MAY-2002 (first entry)  
 XX  
 XX cDNA encoding clone #22045 of lung tumour protein.  
 DE  
 XX Lung tumour; cancer; T cell; immune response stimulator;  
 KW cytostatic; gene; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WC200204514-A2.  
 PN  
 XX 17-JAN-2002.  
 PD  
 XX 10-JUL-2001; 2001WO-US22058.  
 XX  
 XX 11-JUL-2000; 2000US-0614124.  
 PR  
 XX 29-AUG-2000; 2000US-0651563.  
 PF

ER 08-SEP-2000; 2000US-0658824.  
 PR 26-SEP-2000; 2000US-0671325.  
 PR 06-OCT-2000; 2000US-0677419.  
 PR 30-OCT-2000; 2000US-0702705.  
 PR 13-DEC-2000; 2000US-0736457.  
 PR 03-MAY-2001; 2001US-0849626.  
 XX  
 XX (CORI-) CORIYA CORP.  
 PA  
 XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;  
 XX Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;  
 PI Mcnabb A, Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;  
 XX WPI; 2002-164634/21.  
 XX Novel polynucleotide encoding a lung tumour polypeptide useful for  
 PT stimulating and/or expanding T cells specific for a tumour protein -  
 PT  
 XX Example 1; SEQ ID No 233; 223pp; English.  
 PS  
 XX The invention describes an isolated polynucleotide and polypeptide  
 CC useful for stimulating and/or expanding T cells specific for a tumour  
 CC protein for determining the presence of a cancer in a patient. A  
 CC composition containing the polynucleotide and/or polypeptide is useful  
 CC for treating a lung cancer in a patient. The polypeptide is useful for  
 CC removing tumour cells from a biological sample. The polynucleotide is  
 CC also useful as probe or primer to detect the level of mRNA encoding a  
 CC tumour protein. This sequence encodes a lung tumour associated protein  
 CC or protein fragment, described in the method of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 600 BP; 198 A; 106 C; 135 G; 161 T; 0 other;  
 SQ  
 Query Match 10.6%; Score 121; DB 24; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-53;  
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1023 AAGTAAATTTGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCCATT 1082  
 Db 600 AAGTAAATTTGTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCCATT 541  
 QY 1083 CCTATTGTTATCAGGAAGATGACATCATCTTATTTCAGTGGCAAGTTCTTGCCCTTG 1142  
 Db 540 CCTATTGTTATCAGGAAGATGACATCATCTTATTTCAGTGGCAAGTTCTTGCCCTTG 481  
 QY 1143 A 1143  
 Db 480 A 480  
 RESULT 10  
 ACA10555/c  
 ID ACA10555 standard; cDNA; 600 BP.  
 XX  
 XX ACA10555;  
 XX  
 XX 05-JUN-2003 (first entry)  
 XX  
 XX Human lung cancer-associated cDNA, SEQ ID 233.  
 DE  
 XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;  
 KW T cell expansion; CD4; CD8.  
 XX  
 XX Homo sapiens.  
 OS  
 XX US2002197669-A1.  
 PN  
 XX 26-DEC-2002.  
 PD  
 XX 03-MAY-2001; 2001US-0849626.  
 PF

XX 13-DEC-2000; 2000US-0736457.  
PR (BANG/) BANGUR C S.  
PA (FANG/) FANGER G R.  
PA (WANG/) WANG A.  
PA (WANG/) WANG T.  
PA (SWIT/) SWITZER A P.  
PA (MCNE/) MCNEILL P D.  
PA (CLAP/) CLAPPER J D.  
XX  
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;  
PI Clapper JD;  
XX  
XX WPI; 2003-352750/33.  
XX  
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful  
PT for detecting the presence of lung cancer in a patient, and in  
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer -  
XX  
XX Example 1; Page -; 72pp; English.  
XX  
XX The invention relates to a polynucleotide encoding a lung tumour protein,  
CC comprising a sequence selected from any of the 14 sequences  
CC mentioned in the specification, or a sequence (S2) mentioned in  
CC specification, complement of S1, sequences consisting of at least 20  
CC contiguous residues of S1, sequences that hybridise to S1, sequences  
CC having 75%, preferably 90%, identity to S1, or degenerate variants of  
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)  
CC selected from any one of the 4 amino acid sequences mentioned in the  
CC specification, a sequence encoded by the polynucleotide, or sequences  
CC having at least 70%, preferably 90%, identity to a sequence encoded by  
CC the polynucleotide), an expression vector comprising the polynucleotide  
CC operably linked to an expression control sequence, a host cell  
CC transformed or transfected with the vector, an isolated antibody (or its  
CC antigen-binding fragment) that specifically binds to the polypeptide,  
CC detecting the presence of a cancer in a patient, a fusion protein  
CC comprising the polypeptide, an oligonucleotide that hybridises to  
CC S1 under moderately stringent conditions, stimulating and/or expanding T  
CC cells specific for a tumour protein (comprising contacting T cells with  
CC the polynucleotide, protein or antigen-presenting cells, under conditions  
CC and for a time sufficient to permit the stimulation and/or expansion of T  
CC cells; and inhibiting the development of a cancer in a patient (by  
CC incubating CD4+ and/or CD8+ T cells isolated from a patient with the  
CC polynucleotide, protein or antigen presenting cells that express the  
CC polynucleotide, such that T cells proliferate, administering to the  
CC patient an effective amount of the proliferated T cells, and thus  
CC inhibiting the development of a cancer in the patient. The  
CC polynucleotide, protein and cells are useful in a composition for  
CC stimulating an immune response in a patient, and for treating a cancer in  
CC a patient (particularly lung cancer). The oligonucleotide is useful for  
CC determining the presence of a cancer in a patient. The protein and  
CC oligonucleotides are useful in pharmaceutical compositions, e.g.  
CC vaccines. The polynucleotide is also useful as a probe or primer for  
CC nucleic acid hybridisation, and in the design and preparation of  
CC ribozyme molecules for inhibiting expression of tumour polypeptides and  
CC proteins in tumour cells. An amplified portion of the polynucleotide is  
CC useful for isolating a full-length gene from a suitable library.  
CC The present sequence is a cDNA (full length, extended or partial)  
CC isolated from a library derived from lung tumour/cancer cells.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from the USPTO  
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.  
XX  
XX Sequence 600 BP; 198 A; 106 C; 135 G; 161 T; 0 other;  
XX  
XX Query Match 10.6%; Score 121; DB 25; Length 600;  
XX Best Local Similarity 100.0%; Pred. No. 6.6e-53;  
XX Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 600 AAGTAATATTGTAGAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGTGACCAACCCATT 541  
Qy 1083 CCTATTGTTTATCAGGAAGGATGACATCATCTATTCTGCGAAGAGTTCTTSCCCTTG 1142  
Db 540 CCTATTGTTTATCAGGAAGGATGACATCATCTATTCTGCGAAGAGTTCTTSCCCTTG 461  
Qy 1143 A 1143  
Db 480 A 480  
XX  
XX RESULT 11  
XX AEX99506/c  
XX ID AEX99506 standard; cDNA; 600 BP.  
XX AC AEX99506;  
XX  
XX DT 22-MAY-2003 (first entry)  
XX  
XX DE Lung cancer therapy and diagnosis associated cDNA #233.  
XX  
XX KW Lung cancer; cytostatic; vaccine; gene therapy; cancer;  
XX gene; ss.  
XX OS Homo sapiens.  
XX  
XX PN US2002172952-A1.  
XX  
XX PD 21-NOV-2002.  
XX  
XX PF 10-JUL-2001; 2001US-0902941.  
XX  
XX PR 30-JUN-1999; 99US-0346492.  
XX PR 15-OCT-1999; 99US-0419356.  
XX PR 17-DEC-1999; 99US-0466867.  
XX PR 30-DEC-1999; 99US-0476300.  
XX PR 06-MAR-2000; 2000US-0519642.  
XX PR 22-MAR-2000; 2000US-0533077.  
XX PR 10-APR-2000; 2000US-0546259.  
XX PR 27-APR-2000; 2000US-0560406.  
XX PR 05-JUN-2000; 2000US-0589184.  
XX PR 11-JUL-2000; 2000US-0614124.  
XX PR 29-AUG-2000; 2000US-0651563.  
XX PR 08-SEP-2000; 2000US-0658924.  
XX PR 26-SEP-2000; 2000US-0671325.  
XX PR 06-OCT-2000; 2000US-0677419.  
XX PR 30-OCT-2000; 2000US-0702705.  
XX PR 13-DEC-2000; 2000US-0736457.  
XX PR 03-MAY-2001; 2001US-0849626.  
XX  
XX PA (CORI-) CORIXA CORP.  
XX  
XX PI Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;  
XX Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;  
XX  
XX WPI; 2003-328427/31.  
XX  
XX New polynucleotide, useful for preparing a composition for treating or  
XX inhibiting development of cancer, e.g. lung cancer -  
XX  
XX Example 1; SEQ ID NO 233; 82pp; English.  
XX  
XX The invention describes an isolated polynucleotide comprising one of 32  
XX sequences, complement or degenerate variants of them. The polynucleotide  
XX is useful for preparing a composition e.g. a vaccine or for gene therapy,  
XX for treating or inhibiting development of cancer, e.g. lung cancer.  
XX This sequence represents a polynucleotide associated with the  
XX compositions and methods for the therapy and diagnosis of lung cancer.  
XX  
XX Sequence 600 BP; 198 A; 106 C; 135 G; 161 T; 0 other;  
XX  
XX Query Match 10.6%; Score 121; DB 25; Length 600;  
XX Best Local Similarity 100.0%; Pred. No. 6.6e-53;

T	X	14-JUL-1999 (first entry)
E	X	Mouse megsin DNA.
X	X	Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
W	W	human; rat; murine; ss.
W	W	
X	X	Mus musculus.
S	S	
SS	SS	
H	H	Location/Qualifiers
T	T	1..1107
T	T	/tag= a
T	T	/product= "megsin"
T	T	/note= "Partial sequence, no start codon given"
XX	XX	
XX	XX	WO9915652-A1.
NX	NX	
KN	KN	01-APR-1999.
XX	XX	
XX	XX	22-SEP-1998; 98WO-JP04269.
PF	PF	
XX	XX	
PR	PR	22-SEP-1997; 97JP-0275302.
XX	XX	
KA	KA	[KURO/] KUROKAWA K.
FA	FA	(MIYA/) MIYATA T.
XX	XX	
PI	PI	Miyata T;
XX	XX	
DR	DR	WPI; 1999-276983/23.
R	R	P-PSDB; AAY08256.
XX	XX	
PT	PT	Megsin protein expressed specifically in mesangial cells
XX	XX	
PS	PS	Claim 4; Page 72-76; 100pp; Japanese.
CC	CC	This invention describes the isolation of novel megin nucleic acid and
CC	CC	proteins from human, rat and mouse tissue. This protein is expressed
CC	CC	specifically in mesangial cells. The products of the invention are
CC	CC	useful for the treatment and diagnosis of diseases involving mesangial
CC	CC	cells, such as IgA nephropathy.
XX	XX	
SQ	SQ	Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;
		Query Match 2.7%; Score 31; DB 20; Length 1147;
		Best Local Similarity 100.0%; Pred No. 1.4e-05;
		Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	QY	832 CAGTTCGAAGATGAGAAGAATTATGAAATGA 862       796 CAGTTCGAAGATGAGAAGAATTATGAAATGA 826
Db	Db	
		RESULT 14
		AAC55239
ID	ID	AAC55239 standard; DNA; 1147 BP.
XX	XX	
AC	AC	AAC55239;
XX	XX	
DT	DT	30-JAN-2001 (first entry)
DE	DE	
DE	DE	Mouse megsin nucleotide sequence SEQ ID NO:20.
KW	KW	Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
KN	KN	IgA; immunoglobulin A; detection; renal function; renal disorder;
KW	KW	diagnosis; biological sample; blood; urine; ss.
XX	XX	
OS	OS	Mus musculus.
XX	XX	
PN	PN	WO200057189-A1.
XX	XX	
PD	PD	28-SEP-2000.
XX	XX	
FF	FF	17-MAR-2000; 2000WO-JP01646.

```

XX 19-MAR-1999; 99JP-0075305.
XX 28-OCT-1999; 99JP-0306623.
XX (KURO/) KUROKAWA K.
XX (FUSO) FUSO PHARM IND LTD.
XX (MIYA/) MIYATA T.
XX Miyata T;
XX WPI; 2000-511642/58.
XX P-PSDB; AA324151.
XX Evaluating renal function comprises assaying megin protein in
XX biological sample -
XX Disclosure; Page 85-88; 93pp; Japanese.
XX The present invention describes a method for evaluating renal function.
XX The method comprises assaying megin protein in biological sample. Also
XX described are: (1) use of an antimegin protein antibody for diagnosing
XX renal function; and (2) a kit for detecting megin protein comprising:
XX (a) antimegin protein antibody attached to solid magnetic particles;
XX (b) direct or indirect fixing for the antibody to the particles; and
XX (c) a magnet. The process is useful for evaluating renal function and
XX diagnosing renal disorders by assaying megin protein in biological
XX samples (preferably urine or blood). The process is reproducible and
XX gives accurate results. The present sequence encodes the mouse megin
XX protein, which is given in the exemplification of the present invention.
XX Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;
SQ
Query Match 2.7%; Score 31; DB 21; Length 1147;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 832 CAGTTCAGATAGAGAGAGATTATGAAATGA 862
Db 796 CAGTTCAGATAGAGAGAGATTATGAAATGA 826
RESULT 15
AAF82440
ID AAF82440 standard; cDNA; 1147 BP.
XX AC AAF82440;
XX DT 10-JUL-2001 (first entry)
XX DE Mouse megin nucleotide sequence.
XX KW Mouse; megin; mesangial cell proliferative nephritis; nephrotropic;
XX transgenic mouse; glomerular disease; animal model; drug screening; ss.
XX OS Mus musculus.
XX FH Key Location/Qualifiers
XX FT CDS 1..1107
XX FT /*tag= a
XX FT /product= "megsin"
XX FN WO200124628-A1.
XX PD 12-APR-2001.
XX PP 06-OCT-2000; 2000WO-JP06988.
XX PR 06-OCT-1999; 99JP-0285736.
XX PA (KURO/) KUROKAWA K.
XX PA (MIYA/) MIYATA T.
XX PI Miyata T;

```

```

XX WPI; 2001-300136/31.
XX P-PSDB; AA83077.
XX Mouse model for mesangial cell proliferative nephritis for development
XX and screening of new treatments -
XX Disclosure; Page 50-52; 62pp; Japanese.
XX The present sequence encodes mouse megin. The human megin coding
XX sequence may be introduced into a mouse to produce an animal model of
XX mesangial cell proliferative nephritis. The symptoms include
XX enlargement of the mesangial base region, sedimentation of an immune
XX complex and an increase in mesangial cells. The animal model is useful
XX for analysing the pathology of chronic glomerular diseases and for
XX screening compositions for prevention and treatment of the diseases.
XX Highly uniform models can be made easily and in large numbers using
XX this method.
XX Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;
SQ
Query Match 2.7%; Score 31; DB 22; Length 1147;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 832 CAGTTCAGATAGAGAGAGATTATGAAATGA 862
Db 796 CAGTTCAGATAGAGAGAGATTATGAAATGA 826
Search completed: December 14, 2003, 04:48:33
Job time : 381 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2003, 02:33:04 ; Search time 426 Seconds  
(without alignments)  
8917.544 Million cell updates/sec

Title: US-09-508-997A-1

Perfect score: 1143

Sequence: 1 atggcctccctgtgcgcg.....Gcaagttttgtcccttga 1143

Scoring table: OLIGO NUC

Gapop 50.0 , Gapext 60.0

Searched: 2201672 seqs, 1661799599 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1143	100.0	1950	9	US-09-140-719-30
2	1143	100.0	1950	14	US-10-091-442-30
3	399	34.9	399	13	US-10-029-386-21145
4	399	34.9	592	13	US-10-029-386-7445
5	121	10.6	521	13	US-10-029-386-10599
6	121	10.6	600	10	US-09-736-457-233
7	121	10.6	600	10	US-09-902-941-233
8	121	10.6	600	10	US-09-849-626-233
9	121	10.6	600	11	US-09-476-300-233
10	121	10.6	600	13	US-10-113-872-233
11	121	10.6	600	15	US-10-017-754-233
12	119	10.4	119	13	US-10-029-386-24299
13	25	2.2	37	9	US-09-140-719-32
14	25	2.2	37	14	US-10-091-442-32
15	23	2.0	23	9	US-09-140-719-14

16	23	2.0	23	14	US-10-091-442-14	Sequence 14, Appl
17	21	1.8	21	9	US-09-140-719-15	Sequence 15, Appl
18	21	1.8	21	9	US-09-140-719-16	Sequence 16, Appl
19	21	1.8	21	9	US-09-140-719-27	Sequence 27, Appl
20	21	1.8	21	9	US-09-140-719-28	Sequence 28, Appl
21	21	1.8	21	9	US-09-140-719-29	Sequence 29, Appl
22	21	1.8	21	14	US-10-091-442-15	Sequence 15, Appl
23	21	1.8	21	14	US-10-091-442-16	Sequence 16, Appl
24	21	1.8	21	14	US-10-091-442-28	Sequence 27, Appl
25	21	1.8	21	14	US-10-091-442-28	Sequence 28, Appl
26	21	1.8	21	14	US-10-091-442-29	Sequence 29, Appl
27	21	1.8	21	13	US-10-392-219-3	Sequence 3, Appl
28	21	1.8	21	13	US-10-225-810-9	Sequence 9, Appl
29	21	1.8	21	13	US-10-225-810-26	Sequence 26, Appl
30	20	1.7	343	10	US-09-878-574-2608	Sequence 2608, Ap
31	20	1.7	545	13	US-10-027-632-202376	Sequence 202376,
32	20	1.7	545	14	US-10-027-632-202376	Sequence 202376,
33	20	1.7	83450	11	US-09-811-469-3	Sequence 3, Appl
34	20	1.7	83450	13	US-10-370-659-3	Sequence 3, Appl
35	19	1.7	211	9	US-09-864-761-18550	Sequence 18550, A
36	19	1.7	351	10	US-09-783-590-7053	Sequence 7053, Ap
37	19	1.7	448	13	US-09-814-353-14497	Sequence 14497, A
38	19	1.7	470	9	US-09-864-761-1798	Sequence 1798, Ap
39	19	1.7	478	15	US-10-239-422-5	Sequence 5, Appl
40	19	1.7	502	13	US-09-814-353-1763	Sequence 1763, Ap
41	19	1.7	502	13	US-09-814-353-8113	Sequence 8113, Ap
42	19	1.7	613	15	US-10-239-422-7	Sequence 7, Appl
43	19	1.7	686	13	US-10-027-632-134774	Sequence 134774,
44	19	1.7	686	14	US-10-027-632-134774	Sequence 134774,
45	19	1.7	1293	13	US-10-027-632-124492	Sequence 124492,

#### ALIGNMENTS

RESULT 1

US-09-140-719-30

Sequence 30, Application US/09140719

Patent No. US20010026931A1

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSUJIMOTO, No. US20010026931A1

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenji

APPLICANT: ISHIDA, No. US20010026931A1

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAGUCHI, Kozi

APPLICANT: YAMAGUCHI, No. US20010026931A1

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/474,661

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

OPERATING SYSTEM: PC-DOS/MS-DOS

### RESULT, T 3

ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR18.3  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
OTHER INFORMATION: SWISSPROT HIT: 002739, EVALUE 5.00e-24  
OTHER INFORMATION: NT HIT: D88575.1, EVALUE 0.00e+00  
US-10-029-386-7445

Query Match 34.9%; Score 399; DB 13; Length 592;  
Best Local Similarity 100.0%; Pred. No. 3e-203;  
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 ATTGAAACAAACTGACCTTTGAGATCTTAATGGAATGGAAGCAATCCAGGCGAATGACC 804  
DB 146 ATTGAAACAAACTGACCTTTGAGATCTTAATGGAATGGAAGCAATCCAGGCGAATGACC 205  
QY 805 TCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAAGATGAGAGAAATTTATGAATGAA 864  
DB 206 TCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAAGATGAGAGAAATTTATGAATGAA 265  
QY 865 CAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCT 924  
DB 266 CAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCT 325  
QY 925 GGGATTGCTCGGGGGTCTCTGTATATATCAAGATGATGACAAATCTTACATAGAG 984  
DB 326 GGGATTGCTCGGGGGTCTCTGTATATATCAAGATGATGACAAATCTTACATAGAG 385  
QY 985 GTCACTGAGAGGGGACCGAGGCTACTGCTGCACAGGAGTAAATTTAGAAAAGCAA 1044  
DB 386 GTCACTGAGAGGGGACCGAGGCTACTGCTGCACAGGAGTAAATTTAGAAAAGCAA 445  
QY 1045 CTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCAATTCCTATTTGTTATCAGGAAGAT 1104  
DB 446 CTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCAATTCCTATTTGTTATCAGGAAGAT 505  
QY 1105 GACATCATCTTATTCAGTGGCAAGTTCTTGCCCTGA 1143  
DB 506 GACATCATCTTATTCAGTGGCAAGTTCTTGCCCTGA 544

RESULT 5  
US-10-029-386-10599  
Sequence 10599, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: AEWICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 10599  
LENGTH: 521  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR18.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1  
OTHER INFORMATION: EST HUMAN HIT: AW381394.1, EVALUE 1.00e-58  
OTHER INFORMATION: SWISSPROT HIT: P01014, EVALUE 6.00e-07  
OTHER INFORMATION: NT HIT: D88575.1, EVALUE 3.00e-61  
US-10-029-386-10599

Query Match 10.6%; Score 121; DB 13; Length 521;  
Best Local Similarity 100.0%; Pred. No. 5e-54;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR18.3  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
OTHER INFORMATION: SWISSPROT HIT: 002739, EVALUE 5.00e-24  
OTHER INFORMATION: NT HIT: D88575.1, EVALUE 0.00e+00  
US-10-029-386-2145

Query Match 34.9%; Score 399; DB 13; Length 399;  
Best Local Similarity 100.0%; Pred. No. 2.9e-203;  
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 ATTGAAACAAACTGACCTTTGAGATCTTAATGGAATGGAAGCAATCCAGGCGAATGACC 804  
DB 1 ATTGAAACAAACTGACCTTTGAGATCTTAATGGAATGGAAGCAATCCAGGCGAATGACC 60  
QY 805 TCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAAGATGAGAGAAATTTATGAATGAA 864  
DB 61 TCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAAGATGAGAGAAATTTATGAATGAA 120  
QY 865 CAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCT 924  
DB 121 CAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCT 180  
QY 925 GGGATTGCTCGGGGGTCTCTGTATATATCAAGATGATGACAAATCTTACATAGAG 984  
DB 181 GGGATTGCTCGGGGGTCTCTGTATATATCAAGATGATGACAAATCTTACATAGAG 240  
QY 985 GTCACTGAGAGGGGACCGAGGCTACTGCTGCACAGGAGTAAATTTAGAAAAGCAA 1044  
DB 241 GTCACTGAGAGGGGACCGAGGCTACTGCTGCACAGGAGTAAATTTAGAAAAGCAA 300  
QY 1045 CTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCAATTCCTATTTGTTATCAGGAAGAT 1104  
DB 301 CTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCAATTCCTATTTGTTATCAGGAAGAT 360  
QY 1105 GACATCATCTTATTCAGTGGCAAGTTCTTGCCCTGA 1143  
DB 361 GACATCATCTTATTCAGTGGCAAGTTCTTGCCCTGA 399

RESULT 4  
US-10-029-386-7445  
Sequence 7445, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: AEWICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 7445  
LENGTH: 592  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR18.3  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6



QY 335 AGGACTACATGAGTGTCCGGAATAATTATACATGTCACAAAGTGGAGCGAGTTGACTTTA 394  
|  
Db 326 AGGACTACATGAGTGTCCGGAATAATTATACATGTCACAAAGTGGAGCGAGTTGACTTTA 385  
|  
QY 395 CGAATCATTTAGAGACACACTAGACGTAAATTAATAAGTGGTGGTGAATGAACACATG 454  
|  
Db 386 CGAATCATTTAGAGACACACTAGACGTAAATTAATAAGTGGTGGTGAATGAACACATG 445  
|  
QY 455 G 455  
|  
Db 446 G 446

## RESULT 6

US-09-736-457-233/c  
; Sequence 233, Application US/09736457  
; Patent No. US20020168637A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1964  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 233  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-736-457-233

Query Match 10.6%; Score 121; DB 10; Length 600;  
Best Local Similarity 100.0%; Pred. No. 5e-54;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1023 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCAACCCATT 1082  
|  
Db 500 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCAACCCATT 541  
|  
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142  
|  
Db 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481  
|  
QY 1143 A 1143  
|  
Db 480 A 480

## RESULT 7

US-09-902-941-233/c  
; Sequence 233, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 233  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-902-941-233

Query Match 10.6%; Score 121; DB 10; Length 600;  
Best Local Similarity 100.0%; Pred. No. 5e-54;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1023 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCAACCCATT 1082  
|  
Db 500 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCAACCCATT 541  
|  
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142  
|  
Db 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481  
|  
QY 1143 A 1143  
|  
Db 480 A 480

## RESULT 8

US-09-849-626-233/c  
; Sequence 233, Application US/09849626  
; Publication No. US20020197669A1  
; GENERAL INFORMATION:  
; APPLICANT: Bangur, Chaitanya  
; APPLICANT: Fanger, Gary  
; APPLICANT: Wang, Aijun  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Switzer, Anne  
; APPLICANT: McNeill, Patricia  
; APPLICANT: Clapper, Jonathan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C16  
; CURRENT APPLICATION NUMBER: US/09/849,626  
; CURRENT FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 1926  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 233  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-849-626-233

Query Match 10.6%; Score 121; DB 10; Length 600;  
Best Local Similarity 100.0%; Pred. No. 5e-54;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1023 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCAACCCATT 1082  
|  
Db 500 AAGTAATATTGTAGAAAGCAACTCCCTCAGTCCAGCGTGTAGAGCTGACCAACCCATT 541  
|  
QY 1083 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 1142  
|  
Db 540 CCTATTGTTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTG 481  
|  
QY 1143 A 1143  
|  
Db 480 A 480

```

RESULT 9
US-09-476-300-233/c
; Sequence 233, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-476-300-233

Query Match      10.6%; Score 121; DB 11; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 1082
Db 600 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 541

QY 1083 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTG 1142
Db 540 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTG 481

QY 1143 A 1143
Db 480 A 480

RESULT 10
US-10-113-872-233/c
; Sequence 233, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kaicos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-233

Query Match      10.6%; Score 121; DB 13; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 1082
Db 600 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 541

QY 1083 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTG 1142
Db 540 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTG 481

```

```

QY 1143 A 1143
Db 480 A 480

RESULT 11
US-10-017-754-233/c
; Sequence 233, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-233

Query Match      10.6%; Score 121; DB 15; Length 600;
Best Local Similarity 100.0%; Pred. No. 5e-54;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 1082
Db 600 AAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCCAATT 541

QY 1083 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTG 1142
Db 540 CCTATTGTTATCAGGAAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTG 481

QY 1143 A 1143
Db 480 A 480

RESULT 12
US-10-029-386-24299
; Sequence 24299, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24299
; LENGTH: 119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR18.3

```

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1  
 OTHER INFORMATION: SKISPROT HIT: P29308, EVALUE 2.00e-05  
 OTHER INFORMATION: ST\_HUMAN HIT: AW81394.1, EVALUE 3.00e-58  
 OTHER INFORMATION: NT HIT: D88575.1, EVALUE 8.00e-61  
 US-10-029-386-24299

Query Match 10.4%; Score 119; DB 13; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-53;  
 Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 337 GACTACATTGAGTGTCCGAAAAATTATACGATGCCAAAGTGGAGCGAGTTGACTTTACG 396  
 DB 1 GACTACATTGAGTGTCCGAAAAATTATACGATGCCAAAGTGGAGCGAGTTGACTTTACG 60

QY 397 ATCATTTAGAGCACTAGACGTATATTATTAAGTGGGTTGAAATGAACACATG 455  
 DB 61 AATCATTTAGAGCACTAGACGTATATTATTAAGTGGGTTGAAATGAACACATG 119

RESULT 13  
 US-09-140-719-32  
 ; Sequence 32, Application US/09140719  
 ; Patent No. US20010026931A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TSUJIMOTO, Masafumi  
 ; APPLICANT: IWASA, Ruyuki  
 ; APPLICANT: TSUROOKA, No. US20010026931A1uo  
 ; APPLICANT: NAKAZATO, Hiroshi  
 ; APPLICANT: MIURA, Kenju  
 ; APPLICANT: ISHIDA, No. US20010026931A1uhiro  
 ; APPLICANT: KURIHARA, Tatsuya  
 ; APPLICANT: YAMAICHI, Kozo  
 ; APPLICANT: YAMAGUCHI, No. US20010026931A1omi  
 TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: P.O. Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/140,719  
 FILING DATE: 08-AUG-1998  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/474,661  
 FILING DATE: 07-JUN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/091,028  
 FILING DATE: 14-JUL-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 4-212305  
 FILING DATE: 17-JUL-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-067339  
 FILING DATE: 04-MAR-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: McGowan, Malcolm K.  
 REGISTRATION NUMBER: 39,300  
 REFERENCE/DOCKET NUMBER: 001560-247  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 32:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 37 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-09-140-719-32

Query Match 2.2%; Score 25; DB 9; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 0.014;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGCTGCGAAATG 25  
 DB 13 ATGGCTCCCTTGCTGCGAAATG 37

RESULT 14  
 US-10-091-442-32  
 ; Sequence 32, Application US/10091442  
 ; Publication No. US20020164711A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TSUJIMOTO, Masafumi  
 ; APPLICANT: IWASA, Ruyuki  
 ; APPLICANT: TSUROOKA, No. US20020164711A1uo  
 ; APPLICANT: NAKAZATO, Hiroshi  
 ; APPLICANT: MIURA, Kenju  
 ; APPLICANT: ISHIDA, No. US20020164711A1uhiro  
 ; APPLICANT: KURIHARA, Tatsuya  
 ; APPLICANT: YAMAICHI, Kozo  
 ; APPLICANT: YAMAGUCHI, No. US20020164711A1omi  
 TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: P.O. Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/091,442  
 FILING DATE: 07-Mar-2002  
 CLASSIFICATION: <unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/140,719  
 FILING DATE: 08-AUG-1998  
 APPLICATION NUMBER: US 08/474,661  
 FILING DATE: 07-JUN-1995  
 APPLICATION NUMBER: US 08/091,028  
 FILING DATE: 14-JUL-1993  
 APPLICATION NUMBER: JP 4-212305  
 FILING DATE: 17-JUL-1992  
 APPLICATION NUMBER: JP 6-067339  
 FILING DATE: 04-MAR-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: McGowan, Malcolm K.  
 REGISTRATION NUMBER: 39,300  
 REFERENCE/DOCKET NUMBER: 001560-247  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 32:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 37 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single

Mon Dec 15 08:51:17 2003

Query Match 2.0%; Score 23; DB 9; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 CAAGGCGATGACCTCTAAGTAT 813  
 DB 1 CAAGGCGATGACCTCTAAGTAT 23

Search completed: December 14, 2003, 04:42:01  
 Job time : 430 secs

TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
 US-10-091-442-32

Query Match 2.2%; Score 25; DB 14; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 0.014;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCTCCCTGCTGCGAAGT 25  
 DB 13 ATGGCCTCCCTGCTGCGAAGT 37

RESULT 15  
 US-09-140-719-14  
 ; Sequence 14, Application US/09140719  
 ; Patent No. US20010026931A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TSUJIMOTO, Masafumi  
 ; APPLICANT: IWASA, Fuyuki  
 ; APPLICANT: TSUBOUKA, No. US20010026931A1no  
 ; APPLICANT: NAKAZATO, Hiroshi  
 ; APPLICANT: MIURA, Kenju  
 ; APPLICANT: ISHIDA, No. US20010026931A1unhiro  
 ; APPLICANT: KURIHARA, Tatsuya  
 ; APPLICANT: YAMAUCHI, Kozo  
 ; APPLICANT: YAMAGUCHI, No. US20010026931A1omi  
 ; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
 ; NUMBER OF SEQUENCES: 34  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSES: Burns, Doane, Swecker & Mathis  
 ; STREET: P.O. Box 1404  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: United States  
 ; ZIP: 22313-1404  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/140,719  
 ; FILING DATE: 08-AUG-1998  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/474,661  
 ; FILING DATE: 07-JUN-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/091,028  
 ; FILING DATE: 14-JUL-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 4-212305  
 ; FILING DATE: 17-JUL-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 6-067339  
 ; FILING DATE: 04-MAR-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McGowan, Malcolm K.  
 ; REGISTRATION NUMBER: 39,300  
 ; REFERENCE/DOCKET NUMBER: 001560-247  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-6620  
 ; TELEFAX: (703) 836-2021  
 ; INFORMATION FOR SEQ ID NO: 14:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 US-09-140-719-14

Result No.	Score	Query Match	Length	DB ID	Description
c 1	555	48.6	638	9	AW381394 PMO-H7030
2	536	46.9	824	12	B1871335
3	510	44.6	1026	10	BG682538
4	447	39.1	498	13	BX284040
					EX284040 BX284040



```

Db 432 ATATGTAGAAAGCACTCCCTCAGTCACGCTGTTAGAGTCACACCACTCTCTAT 491
Qy 1088 TTGTATCAGGAAGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTGA 1143
Db 452 TTGTATCAGGAAGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCTTGA 547

RESULT 3
EG682538 1026 bp mRNA linear EST 01-MAY-2001
LOCUS 602624811F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749699 5'
DEFINITION mRNA sequence.
ACCESSION EG682538
VERSION 1 (bases 1 to 1026)
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgl.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10603 row: f column: 04
High quality sequence start: 3
High quality sequence stop: 791.

FEATURES
source
1..1026
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749699"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 303 a 221 c 213 g 289 t
ORIGIN
Query Match 44.6%; Score 510; DB 10; Length 1026;
Best Local Similarity 99.8%; Pred. No. 4.5e-262;
Matches 560; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 583 TTCAATCTCCCAAGTGTCTCGGAAGCAGTCGCCATGATGATCAGCAAGCAAGTTC 642
Db 9 TTCAATCTCCCAAGTGTCTCGGAAGCAGTCGCCATGATGATCAGCAAGCAAGTTC 68

Qy 643 AATTTGTCTCTATTGAGGACCCATCAATGAAGATTCTTGAGCTCAGATACAAATGGTGGC 702
Db 69 AATTTGTCTCTATTGAGGACCCATCAATGAAGATTCTTGAGCTCAGATACAAATGGTGGC 128

Qy 703 ATAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 762
Db 129 ATAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 188

Qy 763 TTTCAGATCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 822
Db 189 TTTCAGATCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 248

Qy 823 TTTTTCCTCAGTCAAGATAGAGAAATTTAGAAATGAAGAAACAATATTGAGAGCCCTA 882

```

```

Db 249 TTTTTCCTCAGTCAAGATAGAGAAATTTAGAAATGAAGAAACAATATTGAGAGCCCTA 308
Qy 883 GGGCTGAAGATATCTTTTATGATGAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGT 942
Db 309 GGGCTGAAGATATCTTTTATGATGAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGT 368
Qy 943 GGTCTGTATATATCAAGATGATCCCAATCTTACATAGAGTCTCAGTCCAGGGCACC 1002
Db 369 GGTCTGTATATATCAAGATGATCCCAATCTTACATAGAGTCTCAGTCCAGGGCACC 428
Qy 1003 GAGGCTACTCTGCCACAGGAAGTAATTTGTAGAAAAGCAATCTCCTCAGTCCAGGGT 1062
Db 429 GAGGCTACTCTGCCACAGGAAGTAATTTGTAGAAAAGCAATCTCCTCAGTCCAGGGT 488
Qy 1063 TTTAGAGTCAACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1122
Db 489 TTTAGAGTCAACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 548
Qy 1123 GGCAGAGTTCTTGCCCTTGA 1143
Db 549 GGCAGAGTTCTTGCCCTTGA 569

RESULT 4
EX284040 498 bp mRNA linear EST 05-MAR-2003
LOCUS EX284040 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:998F0410603 ;
DEFINITION IMAGE:4749699, mRNA sequence.
ACCESSION EX284040
VERSION EX284040.1 GI:28848494
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 498)
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
Human UnigeneSet - RZPD3
Unpublished
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPDLib: I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLib No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Neubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
SP6, Primer sequence: ATTTAGTGCACATATAG.

FEATURES
source
1..498
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:998F0410603 ; IMAGE:4749699"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 153 a 101 c 114 g 130 t
ORIGIN

```

Query Match 39.1%; Score 447; DB 13; Length 498;  
 Best Local Similarity 99.8%; Pred. No. 2.5e-228;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 583 TTCATATCCCAAGTCTCTGGAGGCGATGCCATGATCATCAGGAACGGAAGTTC 642  
 Db 1 TTCATATCCCAAGTCTCTGGAGGCGATGCCATGATCATCAGGAACGGAAGTTC 60

QY 643 AATTGCTCTGTTATGAGGACCAATCAATGAAGATCTTGGCTCAGATCAATGTTGTC 702  
 Db 61 AATTGCTCTGTTATGAGGACCAATCAATGAAGATCTTGGCTCAGATCAATGTTGTC 120

QY 703 ATAAACATGATCGTTCTGCTGCTGAGGATGACCTCTCTGAAATTTGAAACAAATGACC 762  
 Db 121 ATAAACATGATCGTTCTGCTGCTGAGGATGACCTCTCTGAAATTTGAAACAAATGACC 180

QY 763 TTTCAGATCTAATGGAATGAGCAATCAAGGGGATGACCTCTAAGTATGTTAGGTA 822  
 Db 181 TTTCAGATCTAATGGAATGAGCAATCAAGGGGATGACCTCTAAGTATGTTAGGTA 240

QY 823 TTTTTCCTCAGTTCAAGATAGAGAAATATGAATGAACAAATATTTGAGAGCCCTA 882  
 Db 241 TTTTTCCTCAGTTCAAGATAGAGAAATATGAATGAACAAATATTTGAGAGCCCTA 300

QY 883 GGCTCAAGAGATCTTTGATGATCAGAGAGATCTCTGGGATGCTTGGGGGT 942  
 Db 301 GGCTCAAGAGATCTTTGATGATCAGAGAGATCTCTCTGGGATGCTTGGGGGT 360

QY 943 CGTCTGTATATCAAGATGATGACAAATCTTACATAGAGTCACTGAGAGGGCACC 1002  
 Db 361 CGTCTGTATATCAAGATGATGACAAATCTTACATAGAGTCACTGAGAGGGCACC 420

QY 1003 GAGCTACTGCTGCCACAGGAGTATATGAGAGAGCAATCTCCCTCAGTCCAGCTG 1062  
 Db 421 GAGCTACTGCTGCCACAGGAGTATATGAGAGAGCAATCTCCCTCAGTCCAGCTG 480

QY -063 TTTAGAGCTGACCCCA 1080  
 Db 481 TTTAGAGCTGACCCCA 498

RESULT 5  
 AW381390/c 638 bp mRNA linear EST 04-FEB-2000  
 LOCUS PMO-HT0302-271099-001-d11 HT0302 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION AW381390  
 ACCESSION AW381390.1 GI:6886049  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 638)  
 HCGP <http://www.ludwig.org.br/ORESTES>.  
 TITLE The FAPESP/LICR Human Cancer Genome Project  
 JOURNAL Unpublished  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PMO&t2=PMO-HT0302-271099-001-d11&t3=1999-10-27&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence start: 30  
 High quality sequence stop: 562.

FEATURES  
 Location/Qualifiers  
 1..638

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0302"  
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 174 a 152 c 131 g 181 t  
 ORIGIN

Query Match 38.1%; Score 435; DB 9; Length 638;  
 Best Local Similarity 99.6%; Pred. No. 7.5e-222;  
 Matches 535; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 84 GTTCTTTTCTCTCTGAGCCTCTTCGCTGCCCTGCTCCGCTTGGCGCTCAAGA 143  
 Db 555 GTTCTTTTCTCTCTGAGCCTCTTCGCTGCCCTGCTCCGCTTGGCGCTCAAGA 496

QY 144 TGACTCCCTCTCTCAGATTGATTAAGTTGCTTCATGTTAAACATCGCTCAGGATATGAAA 203  
 Db 495 TGACTCCCTCTCTCAGATTGATTAAGTTGCTTCATGTTAAACATCGCTCAGGATATGAAA 436

QY 204 CTCTCTTAATAGTCTCAGGCTCCAGTCTCACTGAAAGAGTTTTTCTGATATAA 263  
 Db 435 CTCTCTTAATAGTCTCAGGCTCCAGTCTCACTGAAAGAGTTTTTCTGATATAA 376

QY 264 TGCATCCCAACAGGATTAATGATCTCAGCATTTGTGAATGGCTTTTGTGAAAAAGTGA 323  
 Db 375 TGCATCCCAACAGGATTAATGATCTCAGCATTTGTGAATGGCTTTTGTGAAAAAGTGA 316

QY 324 TGGCTTTTATAAGGACTACATTCAGTGTCCCGAAAAATTTACGATGCCAAAGTGGAGCG 383  
 Db 315 TGGCTTTTATAAGGACTACATTCAGTGTCCCGAAAAATTTACGATGCCAAAGTGGAGCG 256

QY 384 AGTTGACTTTACGATCATTTAGAGACACTAGACTAATTAATAAGTGGTTGAAA 443  
 Db 255 AGTTGACTTTACGATCATTTAGAGACACTAGACTAATTAATAAGTGGTTGAAA 196

QY 444 TCAACACATGCAAAATCAAGACGCTGATGTTGAGCTGGCATTAAGCTCATCTGCTGT 503  
 Db 195 TCAACACATGCAAAATCAAGTACGTTGTTGAGCTGGCATTAAGCTCATCTGCTGT 136

QY 504 AATGGTCTGGTGAATGCTGTGTACTTTCAAAGGCAAGTGGCAATCAGCTTCCCAAGAG 563  
 Db 135 AATGGTCTGGTGAATGCTGTGTACTTTCAAAGGCAAGTGGCAATCAGCTTCCCAAGAG 76

QY 564 CGAAACCAATAAATGCCATTTCAATCTCCAGTGTCTTGGGAAGCAGTCCCAT 620  
 Db 75 CGAAACCAATAAATGCCATTTCAATCTCCAGTGTCTTGGGAAGCAGTCCCAT 19

RESULT 6  
 BG185954 786 bp mRNA linear EST 21-APR-2001  
 LOCUS RST4909 Athensys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BG185954  
 ACCESSION BG185954.1 GI:13707641  
 VERSION BG185954.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 786)  
 Harrington, J. J., Sherif, B., Rundlett, S., Jackson, P. D., Perry, R.,  
 Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,  
 Lerner, D., Costanzo, D., McEligott, K., Booser, S., Mays, R., Smith,  
 E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offendbacher,  
 J., Danzig, J. and Ducar, M.



VERSION	EG319597.1	GI:13129272
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 405)	
AUTHORS	Schlager,J.J., Benjamin,H.R., Ali,K., Levine,C.F., Dodds,A.D., Avery,D.P., Clark,J.H. and Hofmann,C.M.	
TITLE	Human Epidermal Keratinocyte Subtraction Library-Downregulated Transcripts	
JOURNAL	Unpublished	

CONTACT: Schlager, J.J.  
 Molecular Toxicology  
 United States Army Medical Research Institute of Chemical Defense  
 (USAMRICD)  
 3100 Ricketts Point Road; ATTN: MCMR-UV-PA; AFG-EA, MD 21010-5400,  
 USA  
 Tel: 410 436 1940  
 Fax: 410 436 1960  
 Email: John.Schlager@AMRI.MIL.  
 Location/Qualifiers  
 1. 405  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /sex="female"  
 /tissue\_type="skin"  
 /cell\_type="primary keratinocyte"  
 /cell\_line="Clonetics 4075"  
 /dev\_stage="adult"

```

/lab_host="E. coli (ampicillin-resistant)"/
/clone_lib="Human Keratinocyte Subtraction Library-
Downregulated Transcripts"
/notes=Vector: pT-Adv; Library preparation: Sequence
isolated using PCR-select (Clontech) subtraction library
construction (Bhatchenko et al., (1996) Proc. Natl. Acad
Sci. USA 93: 6025-6030) after sulfur mustard exposure of
primary human adult epidermal keratinocytes. Subtraction
library cloned for random sequence selection into pT-Adv
vector. "
BASE COUNT      124 a      81 c      90 g      110 t
ORIGIN
Query Match      35.4%; Score 405; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 8.4e-206;
Matches 405; Conservative 0; Mismatches 0; Gaps 0;
QY      714  CGTCTGCTGCTCGAGATGACCTCTCGAATTGAAACAAACTGACCTTCAGAACT 773
Db      1    CGTCTGCTGCTCGAGATGACCTCTCGAATTGAAACAAACTGACCTTCAGAACT 60
QY      774  AATGGAATGGACCAATCCAAAGCGAATGACCTCTAAGTATGTTGAGGTATTTTTTCCTCA 833
Db      61  AATGGAATGGACCAATCCAAAGCGAATGACCTCTAAGTATGTTGAGGTATTTTTTCCTCA 120
QY      834  GTTCAAGATAGAGACGATATGAAATGAAGAAACAATTTTGGAGAGCCCTAGGGCTGAAGA 893
Db      121  GTTCAAGATAGAGACGATATGAAATGAAGAAACAATTTTGGAGAGCCCTAGGGCTGAAGA 180
QY      894  TATCTTTGATGAATCCAAAGCAGATCTCTCGGATTCGTTCCGGGGGGTCGTGTATAT 953
Db      181  TATCTTTGATGAATCCAAAGCAGATCTCTCGGATTCGTTCCGGGGGGTCGTGTATAT 240

```

234	ATCAAGGATGATCCAAATCTTACATAGAGTCACTGAGAGGGCCGAGGGTACTGC	300	
Db			
Qy	1014	TCCCACGAGGTAATTTGTAGAAAGCAACTCCCTCAGTCCACCGTCTTTAGAGCTGA	1078
Db	301	TGCCACGAGGTAATTTGTAGAAAGCAACTCCCTCAGTCCACCGTCTTTAGAGCTGA	360
Qy	1074	CCACCCATTCCTATTTGTTATTCAGGAAGGATGACATCATCTTATT	1118

Db 361 CCACCATTCTATTGTTATCAGGAGGATGACATCATCTTATT 405

RESULT 8  
B1335844 186 bp mRNA linear EST 30-JUL-2001  
LOCUS B1335844  
DEFINITION 8d119 Human Keratinocyte Subtraction Library- Downregulated  
Transcripts Homo sapiens cDNA similar to serine (or cysteine)  
proteinase inhibitor, clade B (ovalbumin), member 7 (SERPINB7),  
mRNA sequence.  
ACCESSION B1335844  
VERSION B1335844.1 GI:15030578  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 186)  
AUTHORS Schlager, J.J., Benjamin, H.R., Ali, K., Levine, C.F., Poddas, A.D.,  
Avery, D.P., Clark, J.H. and Hofmann, C.M.  
TITLE Human Epidermal Keratinocyte Subtraction Library-Downregulated  
Transcripts  
JOURNAL Unpublished  
COMMENT Contact: Schlager, J.J.  
Molecular Toxicology  
United States Army Medical Research Institute of Chemical Defense  
(USAMRICD)  
3100 Ricketts Point Road; ATTN: MCMR-UV-PA; APG-EA, MD 21010-5400,  
USA  
Tel: 410 436 1940  
Fax: 410 436 1960  
Email: John.Schlager@AMEDD.ARMY.MIL.

FEATURES  
source  
1..186  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="female"  
/tissue\_type="skin"  
/cell\_type="primary keratinocyte"  
/cell\_line="Clonetics 4075"  
/dev\_stages="adult"  
/lab\_host="V5. coli (ampicillin-resistant)"  
/clone\_lib="Human Keratinocyte Subtraction Library-  
Downregulated Transcripts"  
/note="Vector; pT-Adv; Library preparation: Sequence  
isolated using PCR-select (Clontech) subtraction library  
construction (Datchenko et al., (1996) Proc. Natl. Acad.  
Sci. USA 93: 6025-6030) after sulfur mustard exposure of  
primary human adult epidermal keratinocytes. Subtraction  
library cloned for random sequence selection into pT-Adv  
vector."

BASE COUNT 59 a 41 c 41 g 45 t  
ORIGIN  
Query Match 16.3%; Score 186; DB 12; Length 186;  
Best Local Similarity 100.0%; Pred. No. 2.9e-88;  
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 527 ACTTCAGGCGAGTGGCAATCAGCCTTCACGAGGAGGAAACCATATATTCGCAATTCA 586  
DB 1 ACTTCAGGCGAGTGGCAATCAGCCTTCACGAGGAGGAAACCATATATTCGCAATTCA 60  
QY 587 AATCTCCCAAGTCTCTGGGAGGAGTGGCCCATGATCATCAGGAACGGAAGTTCAATT 646  
DB 61 AATCTCCCAAGTCTCTGGGAGGAGTGGCCCATGATCATCAGGAACGGAAGTTCAATT 120  
QY 647 TGTCTGTATTGAGGCCCATCAATGAAGATTTTGAAGCTCAGATACATATGTTGGCATAA 706  
DB 121 TGTCTGTATTGAGGCCCATCAATGAAGATTTTGAAGCTCAGATACATATGTTGGCATAA 180  
QY 707 ACATGT 712

Db 181 ACATGT 186

RESULT 9  
B1247923 627 bp mRNA linear EST 14-NOV-2000  
LOCUS B1247923  
DEFINITION 60185905F1 NIH\_MGC\_58 Homo sapiens cDNA clone IMAGE:4069704 5',  
mRNA sequence.  
ACCESSION B1247923  
VERSION B1247923.1 GI:11163882  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 627)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNL at:  
http://image.llnl.gov  
Plate: JHC912 row: i column: 01  
High quality sequence stop: 385.  
Location/Qualifiers  
1..627  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4069704"  
/tissue\_type="hypernephroma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_58"  
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggccattatggcc); Site\_2: SfiI (ggccattatggcc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGGCATTATGGCC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGGCGGCGGCCGACATG-dT(30)-BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."

BASE COUNT 210 a 129 c 149 g 139 t  
ORIGIN  
Query Match 15.8%; Score 181; DB 10; Length 627;  
Best Local Similarity 100.0%; Pred. No. 2.2e-85;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 426 TAATAGTGGGTTGAAATGAAACACATGCGCAAAATCAAGAACTGATTTGTCGAGGTGG 495  
DB 1 TAATAGTGGGTTGAAATGAAACACATGCGCAAAATCAAGAACTGATTTGTCGAGGTGG 60  
QY 486 CATAGTCTCATCTGCTGTAATGGTGGTGAATCTGTCTTCAAAAGGCAAGTGGCA 545  
DB 61 CATAGTCTCATCTGCTGTAATGGTGGTGAATCTGTCTTCAAAAGGCAAGTGGCA 120  
QY 546 ATCAGCTTCCACAGAGGAGGAAACCATATATTCGCAATTTCCTCCAGTCTCTGG 605  
DB 121 ATCAGCTTCCACAGAGGAGGAAACCATATATTCGCAATTTCCTCCAGTCTCTGG 180  
QY 606 G 606  
DB 181 G 181

RESULT 10	RG622863	1030 bp	mRNA	linear	EST 18-APR-2001
LOCUS	60264768F1 NIH_MGC_79	Human sapiens	cDNA clone IMAGE:4769468	5', mRNA sequence.	
DEFINITION	RG622863	EST			
ACCESSION	RG622863	GI:13674234			
VERSION	RG622863				
KEYWORDS	EST				
SOURCE	Human sapiens (human)				
ORGANISM	Human sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	NIH-MGC	http://mgi.nci.nih.gov/			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: CLONTECH Laboratories, Inc. cDNA library preparation: CLONTECH Laboratories, Inc. cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L10M1634 row: m column: 21 High quality sequence stop: 536.				
FEATURES	BASE COUNT	313 a	283 c	149 g	285 t
ORIGIN	LOCUS	1..1030			
source	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:4769468"				
	/lab_host="DH10B (TI phage-resistant)"				
	/clone_lib="NIH MGC_79"				
	/notes="Organ: placenta; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgcttcggcc); Site 2: SfiI (ggcgcttcggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTCTAGCGGCGGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."				
BASE COUNT	Query Match	10.7%	Score 122;	DB 10;	Length 1030;
ORIGIN	Best Local Similarity	100.0%;	Pred. No. 1.3e-53;		
	Matches 122;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1022	GAAGTAATATTTAGAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACCAACCAT	1081		
DB	1	GAAGTAATATTTAGAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACCAACCAT	60		
QY	1082	TCTATTTGTTATCAGCAAGGATGACATCATCTATTTCAGTGGCAAGTTCTTGGCCCTT	1141		
DB	61	TCTATTTGTTATCAGCAAGGATGACATCATCTATTTCAGTGGCAAGTTCTTGGCCCTT	120		
QY	1142	GA 1143			
DB	121	GA 122			
RESULT 11	D82606	363 bp	mRNA	linear	EST 09-FEB-1996
LOCUS	HUMHBC2926	Human pancreatic islet	Human sapiens	cDNA similar to plasmidogen activator inhibitor 2, mRNA sequence.	
DEFINITION	D82606	EST			
ACCESSION	D82606	GI:1183235			
VERSION	D82606				
KEYWORDS	EST				
SOURCE	Human sapiens (human)				
ORGANISM	Human sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Contact: Jun Takeda				
TITLE	Institute for Molecular and Cellular Regulation, Gunma University				
JOURNAL	3-39-15 Showa-machi, Maebashi Gunma 371, Japan				
COMMENT	Tel: 272-20-8856 Fax: 272-20-8896 Email: jtakeda@b.gunma-u.ac.jp.				
FEATURES	BASE COUNT	104 a	78 c	82 g	10 others
ORIGIN	LOCUS	1..363			
source	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone_lib="Human pancreatic islet"				
	/notes="Vector: Lambda ZAPII; Site 1: Eco RI; Site 2: Xho I; mRNA was prepared from normal adult human islets. cDNA was directionally synthesized from the Xho I in the vector to the EcoRI site. cDNA was size fractionated to remove sequences <1000 bp in size."				
BASE COUNT	Query Match	6.6%;	Score 76;	DB 14;	Length 363;
ORIGIN	Best Local Similarity	100.0%;	Pred. No. 4.7e-29;		
	Matches 76;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	671	TGAAGATTTCTGAGCTGAGTATCATGTGGCATTAACATGATGCTTCTCTGCTGAGA	730		
DB	180	TGAAGATTTCTGAGCTGAGTATCATGTGGCATTAACATGATGCTTCTCTGCTGAGA	239		
QY	731	ATGACCTCTCTGAAAT	746		
DB	240	ATGACCTCTCTGAAAT	255		
RESULT 12	LOCUS	T48573	363 bp	mRNA	linear
T48573	DEFINITION	hbc2926	Human pancreatic islet	Human sapiens	cDNA clone hbc2926
LOCUS	5' end, mRNA sequence.				
ACCESSION	T48573	GI:642773			
VERSION	T48573				
KEYWORDS	EST				
SOURCE	Human sapiens (human)				
ORGANISM	Human sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Human pancreatic islet cDNAs				
TITLE	Bell, G.I. and Takeda, J.				
JOURNAL	Unpublished				
COMMENT	Contact: Bell GI or Takeda J				
FEATURES	BASE COUNT	313 a	283 c	149 g	285 t
ORIGIN	LOCUS	1..363			
source	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db				

```

/clone="hbc2326"
/clone.lib="Human pancreatic islet"
/notes="Vector: Lambda ZAPII; Site 1: Eco RI; Site 2: Xho I; mRNA was prepared from normal adult human islets. cDNA was directionally synthesized from the Xho I in the vector to the EcoRI site. cDNA was size fractionated to remove sequences <1000 bp in size."
BASE COUNT      104 a   78 c   82 g   99 t   10 others
ORIGIN

Query Match      6.6%; Score 76; DB 14; Length 363;
Best Local Similarity 100.0%; Pred. No. 4.7e-29; Mismatches 0; Indels 0; Gaps 0;
Matches 76; Conservative 0;

QY 671 TGAAGATCTTGGAGCTCAGATACAAATGGTGGCATAAACAATGATACGTTCTGCTGCTGAGA 730
      |||||||
Db 180 TGAAGATCTTGGAGCTCAGATACAAATGGTGGCATAAACAATGATACGTTCTGCTGCTGAGA 239

QY 731 ATGACCTCTCGAAAT 746
      |||||||
Db 240 ATGACCTCTCGAAAT 255

RESULT 13
BF243008
LOCUS
DEFINITION
BF243008
ACCESSION
BF243008.1 GI:11156936
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1640)
NIH-MGC http://mgs.nci.nih.gov/;
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CNA Library Preparation: CLONETECH Laboratories, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM985 row: i column: 24
High quality sequence stop: 503.
location/Qualifiers
1. .1640
/organism="Homo sapiens"
/mol type="mRNA"
/db xref="taxon:9606"
/clone="IMAGE:4106207"
/tissue type="from acute myelogenous leukemia"
/lab host="DH10B (TI phage-resistant)"
/clone lib="NIH-MGC 55"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggccattatggcc
sequence: 5'-ATTCTAGAGCGCGCGCGCGCATG-dt(30)BN-3';
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGCGCGCATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT      543 a   377 c   418 g   302 t
ORIGIN

```

```

Query Match      6.3%; Score 72; DB 10; Length 1640;
Best Local Similarity 100.0%; Pred. No. 1.2e-26; Mismatches 0; Indels 0; Gaps 0;
Matches 72; Conservative 0;

QY 1072 GACCACCATTCCTATTGTTATCAGGAAGATGACATCATCTTATTCAGTGGCAAGTT 1131
      |||||||
Db 1 GACCACCATTCCTATTGTTATCAGGAAGATGACATCATCTTATTCAGTGGCAAGTT 60

QY 1132 TCTTGCCTTGA 1143
      |||||||
Db 61 TCTTGCCTTGA 72

RESULT 14
AK014524
LOCUS
DEFINITION
AK014524
ACCESSION
AK014524.1 GI:12852434
VERSION
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, I., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staab, F., Suzuki, R., Tomita, M.,
Wagner, L., Wastio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fleischer, C., Fujita, M., Kamiya, M., Lee, N. H., Lyons, P.,
Hofmann, X., Hume, D. A., Kamei, M., Mazzarelli, J., Mombaez, P., Nordone, P.,
Rincon, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Tayo-Oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawajiri, H., Kotsuki, S.
and Hayashizaki, Y.

```



electrophoresis. Vector DNA was prepared from a derivative of pWP42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 94 a 64 c 75 g 80 t  
ORIGIN

Query Match 2.5%; Score 29; DB 28; Length 313;  
Best Local Similarity 100.0%; Pred. No. 0.0008;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGCTGCAGCAATGCAGA 29  
|||  
Db 250 ATGGCTCCCTTGCTGCAGCAATGCAGA 222  
|||

Search completed: December 14, 2003, 06:48:14  
Job time : 2784 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 13, 2003, 23:52:22 ; Search time 4370 Seconds  
(without alignments)  
10700.161 Million cell updates/sec  
Title: US-09-508-997A-1  
Perfect score: 1143  
Sequence: 1 atggcctccctgtgcagc.....gcaaggttttgccttga 1143

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2889711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match C%  
Maximum Match 100%  
Listing first 45 summaries

Database :					GenEmbl :				
1:	gb	ba	*		1:	gb	ba	*	
2:	gb	htg	*		2:	gb	htg	*	
3:	gb	in	*		3:	gb	in	*	
4:	gb	ov	*		4:	gb	ov	*	
5:	gb	ov	*		5:	gb	ov	*	
6:	gb	pat	*		6:	gb	pat	*	
7:	gb	ph	*		7:	gb	ph	*	
8:	gb	pl	*		8:	gb	pl	*	
9:	gb	pr	*		9:	gb	pr	*	
10:	gb	ro	*		10:	gb	ro	*	
11:	gb	sts	*		11:	gb	sts	*	
12:	gb	sy	*		12:	gb	sy	*	
13:	gb	un	*		13:	gb	un	*	
14:	gb	vi	*		14:	gb	vi	*	
15:	em	ba	*		15:	em	ba	*	
16:	em	fun	*		16:	em	fun	*	
17:	em	hum	*		17:	em	hum	*	
18:	em	in	*		18:	em	in	*	
19:	em	mu	*		19:	em	mu	*	
20:	em	or	*		20:	em	or	*	
21:	em	ov	*		21:	em	ov	*	
22:	em	pat	*		22:	em	pat	*	
23:	em	ph	*		23:	em	ph	*	
24:	em	pl	*		24:	em	pl	*	
25:	em	ro	*		25:	em	ro	*	
26:	em	ro	*		26:	em	ro	*	
27:	em	sts	*		27:	em	sts	*	
28:	em	un	*		28:	em	un	*	
29:	em	vi	*		29:	em	vi	*	
30:	em	htg	hum	*	30:	em	htg	hum	*
31:	em	htg	inv	*	31:	em	htg	inv	*
32:	em	htg	other	*	32:	em	htg	other	*
33:	em	htg	mus	*	33:	em	htg	mus	*
34:	em	htg	pln	*	34:	em	htg	pln	*
35:	em	htg	rod	*	35:	em	htg	rod	*
36:	em	htg	mam	*	36:	em	htg	mam	*
37:	em	htg	vir	*	37:	em	htg	vir	*
38:	em	sy	*		38:	em	sy	*	
39:	em	htgo	hum	*	39:	em	htgo	hum	*
40:	em	htgo	mus	*	40:	em	htgo	mus	*
41:	em	htgo	other	*	41:	em	htgo	other	*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1143	100.0	1143	6	BD094240	BD094240 Model ani
2	1143	100.0	1950	6	AR052445	AR052445 Sequence
3	1143	100.0	1950	6	AR082433	AR082433 Sequence
4	1143	100.0	1950	6	BD08396	BD08396 cDNA encodi
5	1143	100.0	1950	9	D88575	D88575 Homo sapien
6	1143	100.0	2249	9	AF027866	AF027866 Homo sapi
7	781.4	68.4	1386	10	AF105328	AF105328 Mus muscu
8	755.8	66.1	1143	10	AF105329	AF105329 Rattus no
9	755.8	65.1	1229	6	BD096261	BD096261 Method fo
10	753	65.9	1229	6	BD094241	BD094241 Model ani
11	748.6	65.5	1147	6	BD094242	BD094242 Model ani
12	399.4	34.9	157284	9	AC072051	AC072051 Homo sapi
13	399.4	34.9	189092	2	AC027524	AC027524 Homo sapi
14	304	26.6	316	6	AR270261	AR270261 Sequence
15	279.4	24.4	121141	10	AC125314	AC125314 Mus muscu
16	260.2	22.8	51602	2	AC129261	AC129261 Rattus no
17	260.2	22.8	199669	10	AC103453	AC103453 Rattus no
18	260.2	22.8	219071	2	AC133259	AC133259 Rattus no
19	254	22.2	254	6	AR270095	AR270095 Sequence
20	240.6	21.0	1179	9	AF419953	AF419953 Homo sapi
21	240.6	21.0	1179	9	AF419954	AF419954 Homo sapi
22	240.6	21.0	1179	9	AF419955	AF419955 Homo sapi
23	237.8	20.8	1200	6	AX301894	AX301894 Sequence
24	237.4	20.8	1663	10	BC010313	BC010313 Mus muscu
25	225.6	19.7	1512	6	A21239	A21239 H sapiens B
26	225.6	19.7	1512	6	114025	114025 Sequence 3
27	222	19.4	1482	6	A21240	A21240 H sapiens B
28	222	19.4	1482	6	114026	114026 Sequence 5
29	220.6	19.3	1279	6	AX677161	AX677161 Sequence
30	220.6	19.3	1279	9	AF169949	AF169949 Homo sapi
31	220.6	19.3	1308	6	AF176431	AF176431 Sequence
32	220.6	19.3	1308	6	AR220500	AR220500 Sequence
33	220.6	19.3	1308	6	AR255494	AR255494 Sequence
34	220.6	19.3	1308	6	AR281064	AR281064 Sequence
35	220.6	19.3	1308	6	AX365716	AX365716 Sequence
36	219	19.2	1279	6	AX092547	AX092547 Sequence
37	219	19.2	2654	9	HSP113122	AJ001697 Homo sapi
38	219	19.2	3443	9	HSP113711	AJ001696 Homo sapi
39	214.2	18.7	1327	4	D55670	D55670 Bos taurus
40	211.6	18.5	1173	9	AY245781	AY245781 Homo sapi
41	211.6	18.5	1173	9	BT006748	BT006748 Homo sapi
42	211.6	18.5	1173	12	BT007953	BT007953 Synthetic
43	211.6	18.5	1245	6	AX030195	AX030195 Sequence
44	211.6	18.5	1279	9	HSU19556	U19556 Human squam
45	211.6	18.5	1711	6	AR019398	AR019398 Sequence

ALIGNMENTS

RESULT 1  
BD094240  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

BD094240  
Model animals of nephropathy proliferating mesangial cells.  
BD094240  
BD094240.1 GI:22639828  
WO 014628-A/1.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1143)  
Miyata, T.  
Model animals of nephropathy proliferating mesangial cells  
Patent: WO 014628-A 1 12-APR-2001;

COMMENT TOSHIO MIYATA, KIYOSHI KUOKAWA  
OS Homo sapiens (human)  
PN WC 0124628-A/1  
PD 12-APR-2001  
PR 06-OCT-2000 WO 2000P006988  
PF 06-OCT-1999 JP 99P 285736  
PI TOSHIO MIYATA  
PC A01K67/027,A61P13/12,A61K45/00,C12N15/12,C12N15/85,G01N33/15,  
CC G01N33/50  
CH  
EC  
PT

FEATURES  
source  
Key CDS Location/Qualifiers  
CDs (1) (1140).  
1..1143 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 336 a 230 c 261 g 316 t  
ORIGIN

Query Match 100.0%; Score 1143; DB 6; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 9.9e-312;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGCTCCCTGCTGCAGCAAAATGCGAGATTTTGGTTCAACCTGTTTCAGAGAGATGGAAT 60  
Db 1 ATGGCTCCCTGCTGCAGCAAAATGCGAGATTTTGGTTCAACCTGTTTCAGAGAGATGGAAT 60  
Qy 61 GACAATCAAGGAATGGAATGTTCTTTCTCTCTGAGGCTCTTCGCTGCCCTGGCC 120  
Db 61 GACAATCAAGGAATGGAATGTTCTTTCTCTCTCTGAGGCTCTTCGCTGCCCTGGCC 120  
Qy 121 CTGTCGGCTGGGGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGCTTCATGTT 180  
Db 121 CTGTCGGCTGGGGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGCTTCATGTT 180  
Qy 181 AACACTGCTCAGGATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240  
Db 181 AACACTGCTCAGGATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240  
Qy 241 AAAAGAGTTTTTCTGATATAAATGATCCCAAGGATTTATGATCTCAGATTGGAAT 300  
Db 241 AAAAGAGTTTTTCTGATATAAATGATCCCAAGGATTTATGATCTCAGATTGGAAT 300  
Qy 301 GGGCTTTTCTGAAAAGTGTATGCTTTTATAGACTACATTTAGTGTGCGGAAAAA 360  
Db 301 GGGCTTTTCTGAAAAGTGTATGCTTTTATAGACTACATTTAGTGTGCGGAAAAA 360  
Qy 361 TTATACGTCGCAAGTGGAGGAGTTGACTTTTACGATCATTTAGAGACACTAGACCT 420  
Db 361 TTATACGTCGCAAGTGGAGGAGTTGACTTTTACGATCATTTAGAGACACTAGACCT 420  
Qy 421 AATAATTAAAGTGGTTGAAATGAAACACATGGCAAAATCAAGAACGATGTTGTTGAA 480  
Db 421 AATAATTAAAGTGGTTGAAATGAAACACATGGCAAAATCAAGAACGATGTTGTTGAA 480  
Qy 481 GGTGGCTATAGCTCATCTGCTGTAATGCTGCTGTGATGCTGTGACTTCAAGAGCAAG 540  
Db 481 GGTGGCTATAGCTCATCTGCTGTAATGCTGCTGTGATGCTGTGACTTCAAGAGCAAG 540  
Qy 541 TGGCAATCAGCCCTTCAACAGAGCGAAACATTAATTTGCCATTTTCAATCTCCCAAGTC 600  
Db 541 TGGCAATCAGCCCTTCAACAGAGCGAAACATTAATTTGCCATTTTCAATCTCCCAAGTC 600  
Qy 601 TCTGGAGGAGCTGCGCATGATGCAATCAGGAACCGAAGTTCAATTTGCTGTTATTTAG 660  
Db 601 TCTGGAGGAGCTGCGCATGATGCAATCAGGAACCGAAGTTCAATTTGCTGTTATTTAG 660  
Qy 661 GACCCATCAATGAGATTTCTTGTAGCTCAGATCATGTTGGCATTAACATGACTGTTCTG 720  
Db 661 GACCCATCAATGAGATTTCTTGTAGCTCAGATCATGTTGGCATTAACATGACTGTTCTG 720  
Qy 721 CTGCTGAGAAATGACCTCTCTGAAATTTGAAAAACAACTGACCTTTTCAAGATTAATGAA 780

Db 721 CTGCTGAGAAATGACCTCTCTGAAATTTGAAAAACAACTGACCTTTTCAAGATTAATGAA 780  
Qy 781 TGGACCAATCCCAAGCGAATGACCTCTAAAGTATCTTGAAGTATTTTTCCTCAGTTCAAG 840  
Db 781 TGGACCAATCCCAAGCGAATGACCTCTAAAGTATCTTGAAGTATTTTTCCTCAGTTCAAG 840  
Qy 841 ATAGAGAGAAATTTGAATGAAACAAATTTTGAAGCCCTTAGGGCTGAAAGATATCTTT 900  
Db 841 ATAGAGAGAAATTTGAATGAAACAAATTTTGAAGCCCTTAGGGCTGAAAGATATCTTT 900  
Qy 901 GATGATCCAAAGCAGATCTCTCTGGGATTTCTCGGGGGGCTGCTCTGTATATATCAAGG 960  
Db 901 GATGATCCAAAGCAGATCTCTCTGGGATTTCTCGGGGGGCTGCTCTGTATATATCAAGG 960  
Qy 961 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGACCCGAGGCTACTGTCGCACA 1020  
Db 961 ATGATGCAAAATCTTACATAGAGTCACTGAGGAGGACCCGAGGCTACTGTCGCACA 1020  
Qy 1021 GGAAGTAATTTGTAGAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080  
Db 1021 GGAAGTAATTTGTAGAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080  
Qy 1081 TTCTATTTGTTATCAGGAAGATGACATPCATTTATTCAGTGGCAAAAGTTTCTTGCCCT 1140  
Db 1081 TTCTATTTGTTATCAGGAAGATGACATPCATTTATTCAGTGGCAAAAGTTTCTTGCCCT 1140  
Qy 1141 TGA 1143  
Db 1141 TGA 1143

RESULT 2  
AR052445 AR052445 1950 bp DNA linear PAT 29-SEP-1999  
LOCUS Sequence 30 from patent US 5831030.  
DEFINITION AR052445  
ACCESSION AR052445  
VERSION AR052445.1 GI:5975909  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1950)  
AUTHORS Tsujimoto, M., Iwasa, F., Tsuruoka, N., Nakazato, H., Miura, K.,  
Ishida, N., Kurihara, T., Yamaichi, K. and Yamaguchi, N.  
TITLE Antibodies specific for megakaryocyte differentiation factor  
JOURNAL Patent: US 5831030-A 30 03-NOV-1996;  
FEATURES Location/Qualifiers  
source  
1..1950  
/organism="unknown"

BASE COUNT 570 a 384 c 407 g 589 t  
ORIGIN  
Query Match 100.0%; Score 1143; DB 6; Length 1950;  
Best Local Similarity 100.0%; Pred. No. 1e-311;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGCTCCCTGCTGCAGCAAAATGCGAGATTTTGGTTCAACCTGTTTCAGAGAGATGGAAT 60  
Db 74 ATGGCTCCCTGCTGCAGCAAAATGCGAGATTTTGGTTCAACCTGTTTCAGAGAGATGGAAT 133  
Qy 61 GACAATCAAGGAATGGAATGTTCTTTCTCTCTGAGGCTCTTCGCTGCCCTGGCC 120  
Db 134 GACAATCAAGGAATGGAATGTTCTTTCTCTCTGAGGCTCTTCGCTGCCCTGGCC 193  
Qy 121 CTGTCGGCTGGGGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGCTTCATGTT 180  
Db 194 CTGTCGGCTGGGGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGCTTCATGTT 253  
Qy 181 AACACTGCTCAGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240  
Db 254 AACACTGCTCAGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 313



QY 241 AAAAAAGTTTTTTCTGATATAAATGATCCACAGGATTAATGATCTCAGCATTTGTAAT 300  
Db |||||  
QY 314 AAAAAAGTTTTTTCTGATATAAATGATCCACAGGATTAATGATCTCAGCATTTGTAAT 373  
Db |||||  
QY 301 GGGCTTTTCTGAAAAAGTGTATGCTTTTATAGGACTACATTTGAGTGTGCGGAAAAA 360  
Db |||||  
QY 374 GGGCTTTTCTGAAAAAGTGTATGCTTTTATAGGACTACATTTGAGTGTGCGGAAAAA 433  
Db |||||  
QY 361 TTATAGATGCCAAAGTGGAGCGAGTGTGCTTTTACGATCAATTTAGAGACACTAGACGT 420  
Db |||||  
QY 434 TTATAGATGCCAAAGTGGAGCGAGTGTGCTTTTACGATCAATTTAGAGACACTAGACGT 493  
Db |||||  
QY 421 AATATTAAATAGTGGTGTAAATGAAACACATGCGCAAAATCAAGAACGCTGATGTA 480  
Db |||||  
QY 494 AATATTAAATAGTGGTGTAAATGAAACACATGCGCAAAATCAAGAACGCTGATGTA 553  
Db |||||  
QY 481 GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGATGCTGTGATGCTGTGATGCTGTG 540  
Db |||||  
QY 554 GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGATGCTGTGATGCTGTGATGCTGTG 613  
Db |||||  
QY 541 TGGCAATCAGCCCTTCCACAGAGCGAATCAATTAATTTGCCAATTTCAAAATCTCCCAAGTGC 600  
Db |||||  
QY 614 TGGCAATCAGCCCTTCCACAGAGCGAATCAATTAATTTGCCAATTTCAAAATCTCCCAAGTGC 673  
Db |||||  
QY 601 TCTGGGAAGGAGTGGCGCATGATGATCAGGAAAGGAGTTCAAATTTCTGTTTATGAG 660  
Db |||||  
QY 674 TCTGGGAAGGAGTGGCGCATGATGATCAGGAAAGGAGTTCAAATTTCTGTTTATGAG 733  
Db |||||  
QY 661 GACCCCTCAATCAAGATTCCTGAGCTCAGATACATGCTGGGCHATAAATGATGCTGCTG 720  
Db |||||  
QY 734 GACCCCTCAATCAAGATTCCTGAGCTCAGATACATGCTGGGCHATAAATGATGCTGCTG 793  
Db |||||  
QY 721 CTGCTCAGATGAGCTCTCTGAAATTTGAAACCAAACTGACCTTTCAAGATTAATGGA 780  
Db |||||  
QY 754 CTGCTCAGATGAGCTCTCTGAAATTTGAAACCAAACTGACCTTTCAAGATTAATGGA 853  
Db |||||  
QY 781 TGGACCAATCCAAAGCGGAATGACCTCTAAATGATGTTGAGGTATTTTCTCAGTTCAG 840  
Db |||||  
QY 854 TGGACCAATCCAAAGCGGAATGACCTCTAAATGATGTTGAGGTATTTTCTCAGTTCAG 913  
Db |||||  
QY 841 ATAGAGAGAAATTAAGAAATGAAACAAATTTAGAGCCCTAGGCGTGAAGATATCTTT 900  
Db |||||  
QY 914 ATAGAGAGAAATTAAGAAATGAAACAAATTTAGAGCCCTAGGCGTGAAGATATCTTT 973  
Db |||||  
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTGGGGGCTGCTGCTGATATATCAAG 960  
Db |||||  
QY 974 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTGGGGGCTGCTGCTGATATATCAAG 1033  
Db |||||  
QY 961 ATGATGCAAAATCTTACATAGAGGTCACTGAGGAGGACCGAGGCTACTGCTGCCACA 1020  
Db |||||  
QY 1034 ATGATGCAAAATCTTACATAGAGGTCACTGAGGAGGACCGAGGCTACTGCTGCCACA 1093  
Db |||||  
QY 1021 GGAAGTAATTTGTAGAAAGCACTCCCTCAGTCCAGCTGTTTAGAGCTGACACCCA 1080  
Db |||||  
QY 1094 GGAAGTAATTTGTAGAAAGCACTCCCTCAGTCCAGCTGTTTAGAGCTGACACCCA 1153  
Db |||||  
QY 1081 TTCTATTGTTTATCAGAGGATGACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCCT 1140  
Db |||||  
QY 1154 TTCTATTGTTTATCAGAGGATGACATCATCTTATTCAGTGGCAAAAGTTTCTTGCCCT 1213  
Db |||||  
QY 1141 TGA 1143  
Db |||||  
QY 1214 TGA 1216

RESULT 3  
AR082433  
LOCUS 1950 bp DNA linear PAT 31-AUG-2000  
DEFINITION Sequence 30 from patent US 5972886.  
ACCESSION AR082433  
VERSION AR082433.1 GI:10009159  
KEYWORDS Unknown.

ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1950)  
AUTHORS Tsujimoto, M., Iwasa, F., Tsuruoka, N., Nakazato, H., Miura, K.,  
Ishida, N., Kurihara, T., Yamauchi, K. and Yamaguchi, N.  
TITLE Megakaryocyte differentiation factor  
JOURNAL Patent: US 5972886-A 30 26-OCT-1999;  
FEATURES Location/Qualifiers  
source  
1..1950  
/organism="unknown"  
BASE COUNT 570 a 384 c 407 g 589 t  
ORIGIN  
Query Match 100.0%; Score 1143; DB 6; Length 1950;  
Best Local Similarity 100.0%; Pred. No. 1e-311;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCCCTCCCTTCTGTCGACGAAATGACAGAGTTTTCCTTCAACCTGTTTCAGAGAGATGGAT 60  
Db 74 ATGGCCCTCCCTTCTGTCGACGAAATGACAGAGTTTTCCTTCAACCTGTTTCAGAGAGATGGAT 133  
QY 61 GACAAATCAGGAAATGGAATATGTTCTTCTCTCTGAGCCCTCTTCGCTGCCCTGCC 120  
Db 134 GACAAATCAGGAAATGGAATATGTTCTTCTCTCTGAGCCCTCTTCGCTGCCCTGCC 193  
QY 121 CTGCTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATTAAGTTGCTTCATGTT 180  
Db 194 CTGCTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATTAAGTTGCTTCATGTT 253  
QY 181 AACACTCCCTCAGGATATGGAACCTCTTCAATAGTCAAGTCAAGGCTCCAGTCTCAACTG 240  
Db 254 AACACTCCCTCAGGATATGGAACCTCTTCAATAGTCAAGTCAAGGCTCCAGTCTCAACTG 313  
QY 241 AAAAGAGTTTCTGATATTAATGCAATGCCACAGAGTTATGATCTCAGCATTTGTAAT 300  
Db 314 AAAAGAGTTTCTGATATTAATGCAATGCCACAGAGTTATGATCTCAGCATTTGTAAT 373  
QY 301 GGGCTTTTCTGAAAAAGTGTATGCTTTTATAGGACTACATTTGAGTGTGCGGAAAAA 360  
Db 374 GGGCTTTTCTGAAAAAGTGTATGCTTTTATAGGACTACATTTGAGTGTGCGGAAAAA 433  
QY 361 TTATAGATGCCAAAGTGGAGCGAGTGTGCTTTTACGATCAATTTAGAGACACTAGACGT 420  
Db 434 TTATAGATGCCAAAGTGGAGCGAGTGTGCTTTTACGATCAATTTAGAGACACTAGACGT 493  
QY 421 AATATTAAATAGTGGTGTAAATGAAACACATGCGCAAAATCAAGAACGCTGATGTA 480  
Db 494 AATATTAAATAGTGGTGTAAATGAAACACATGCGCAAAATCAAGAACGCTGATGTA 553  
QY 481 GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGATGCTGTGATGCTGTGATGCTGTG 540  
Db 554 GGTGGCATTAAGCTCATCTGCTGTAAATGCTGTGATGCTGTGATGCTGTGATGCTGTG 613  
QY 541 TGGCAATCAGCCCTTCCACAGAGCGAATCAATTAATTTGCCAATTTCAAAATCTCCCAAGTGC 600  
Db 614 TGGCAATCAGCCCTTCCACAGAGCGAATCAATTAATTTGCCAATTTCAAAATCTCCCAAGTGC 673  
QY 601 TCTGGGAAGGAGTGGCGCATGATGATCAGGAAAGGAGTTCAAATTTCTGTTTATGAG 660  
Db 674 TCTGGGAAGGAGTGGCGCATGATGATCAGGAAAGGAGTTCAAATTTCTGTTTATGAG 733  
QY 661 GACCCCTCAATCAAGATTCCTGAGCTCAGATACATGCTGGGCHATAAATGATGCTGCTG 720  
Db 734 GACCCCTCAATCAAGATTCCTGAGCTCAGATACATGCTGGGCHATAAATGATGCTGCTG 793  
QY 721 CTGCTCAGATGAGCTCTCTGAAATTTGAAACCAAACTGACCTTTCAAGATTAATGGA 780  
Db 794 CTGCTCAGATGAGCTCTCTGAAATTTGAAACCAAACTGACCTTTCAAGATTAATGGA 853  
QY 781 TGGACCAATCCAAAGCGGAATGACCTCTAAATGATGTTGAGGTATTTTCTCAGTTCAG 840  
Db 854 TGGACCAATCCAAAGCGGAATGACCTCTAAATGATGTTGAGGTATTTTCTCAGTTCAG 913

```
QY 841 ATAGAGAGAAATTATGAAATGAAACAAATATTGAGAGCCCTAGGCTGAAAGATATCTTT 903
DB 914 ATAGAGAGAAATTATGAAATGAAACAAATATTGAGAGCCCTAGGCTGAAAGATATCTTT 973
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGG 960
DB 974 GATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGG 1033
QY 961 ATGATGCACAAATCTTACATAGAGGTCACTGAGAGGGGACCGAGGCTACTGTCGCCACA 1020
DB 1034 ATGATGCACAAATCTTACATAGAGGTCACTGAGAGGGGACCGAGGCTACTGTCGCCACA 1093
QY 1021 GGAAGTAATTTGTAAGAAAGCACTCCCTCAGTCCAGCTGTTTAGAGCTGACACCCA 1080
DB 1094 GGAAGTAATTTGTAAGAAAGCACTCCCTCAGTCCAGCTGTTTAGAGCTGACACCCA 1153
QY 1081 TTCCTATTCTTATCAGGAAGGATGACATCATCTTATTCAGTGGGCAAGTTCTTGCCCT 1140
DB 1154 TTCCTATTCTTATCAGGAAGGATGACATCATCTTATTCAGTGGGCAAGTTCTTGCCCT 1213
QY 1141 TGA 1143
DB 1214 TGA 1216

RESULT 4
E08396 1950 bp RNA linear PAT 29-SEP-1997
LOCUS cDNA encoding human megakaryocyte growth differentiating factor.
DEFINITION E08396
ACCESSION E08396
VERSION E08396.1 GI:2176513
KEYWORDS JP 1994313000-A/20.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tsujimoto,M., Kurihara,T., Ishida,N., Iwasa,F., Nakazato,H.,
Yamaichi,H., Miura,T., Tsuruoka,N. and Yamaguchi,M.
MEGAKARYOCYTE-PROLIFERATING AND DIFFERENTIATING FACTOR
Patent: JP 1994313000-A 20 NOV-1994;
SUNTORY LTD
OS Homo sapiens (human)
EN JP 1994313000-A/20
PD 08-NOV-1994
PF 16-JUL-1993 JP 1993197752
PR 17-JUL-1992 JP 92P 212305, 04-MAR-1993 JP 92P 67339 PI
TSUJIMOTO MASAFUMI, KURIHARA TATSUYA, ISHIDA NOBUHIRO, PI IWASA
FUYUKI
PI NAKAZATO HIROSHI, YAMAICHI HIROZO, MIURA TAKEHISA, PI
TSURUOKA NOBUO,
PI YAMAGUCHI MARE
PC C07K15/14,A61K37/02,C12N5/10,C12N15/19,C12P21/02,C12P21/02,
C12R1:91;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
PH Key Location/Qualifiers
FT 1..1950
FT source /organism='Homo sapiens'
FT /cell_line='A431 (human epidermoid FT
FT carcinoma)'
FT 1..73
FT 5'UTR
FT CDS 74..1216
FT /product='Human megakaryocyte growth FT
FT differentiating
FT factor'
FT 1217..1950
FT 3'UTR
FT polyA signal 1933..1938.
FT Location/Qualifiers
FT 1..1950
FT source
```

```
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
BASE COUNT 570 a 384 c 407 g 589 t
ORIGIN
Query Match 100.0%; Score 1143; DB 6; Length 1950;
Best Local Similarity 100.0%; Pred. No. 1e-311;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCTCCCTTCTGCTGAGGAAATGAGAGTCTTTCCTTCAACCTGTTGAGAGATGGAT 60
DB 74 ATGCGCTCCCTTCTGCTGAGGAAATGAGAGTCTTTCCTTCAACCTGTTGAGAGATGGAT 133
QY 61 GACATCAAGGAATGGAAATGTTCTTTCTCTCTGAGCCTCTTCGCTCCCTCGCC 120
DB 134 GACATCAAGGAATGGAAATGTTCTTTCTCTCTGAGCCTCTTCGCTCCCTCGCC 193
QY 121 CTGCTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTCAAGATGATTAAGTTGCTTCACTT 180
DB 194 CTGCTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTCAAGATGATTAAGTTGCTTCACTT 253
QY 181 AACACTGCTCAGGATATGGAACCTCTTCTAATAGTCACTGAGGCTCCAGTCTCAACTG 240
DB 254 AACACTGCTCAGGATATGGAACCTCTTCTAATAGTCACTGAGGCTCCAGTCTCAACTG 313
QY 241 AAAGAGTTTTTCTGATATAATGCAATCCCAAGGATTTATGATCTCAGCATTTCTGAAT 300
DB 314 AAAGAGTTTTTCTGATATAATGCAATCCCAAGGATTTATGATCTCAGCATTTCTGAAT 373
QY 301 GGGCTTTTTGCTGAAAGAGTATGGCTTTTCATAGGACTACATTGAGTGTCCGAAAAA 360
DB 374 GGGCTTTTTGCTGAAAGAGTATGGCTTTTCATAGGACTACATTGAGTGTCCGAAAAA 433
QY 361 TTATACGATGCCAAAGTGAAGGAGTTCATTTTCAAGATCAATTTAGAGACACTAGAGGT 420
DB 434 TTATACGATGCCAAAGTGAAGGAGTTCATTTTCAAGATCAATTTAGAGACACTAGAGGT 493
QY 421 ATATATTAATAGTGGTTCGAAATGAAACACATGCGCAAAATCAAGAACGTGATGTTGTA 480
DB 494 ATATATTAATAGTGGTTCGAAATGAAACACATGCGCAAAATCAAGAACGTGATGTTGTA 553
QY 481 GTGGCATAAGCTCATCTCTCTTAATGCTGCTGTAATGCTGTGTGTAATGCTGTGTAATG 540
DB 554 GTGGCATAAGCTCATCTCTCTTAATGCTGCTGTAATGCTGTGTGTAATGCTGTGTAATG 613
QY 541 TGGCAATCAGCTTCCAGAGAGGACCAACCTTAATTTGCCATTTTCAATCTCCAGTGC 600
DB 614 TGGCAATCAGCTTCCAGAGAGGACCAACCTTAATTTGCCATTTTCAATCTCCAGTGC 673
QY 601 TCTGGAGAGGAGTGGCCATGATGATCAGGAACGGAGTTCAATTTGCTGTTATGAG 660
DB 674 TCTGGAGAGGAGTGGCCATGATGATCAGGAACGGAGTTCAATTTGCTGTTATGAG 733
QY 661 GACCCATCAATGAAGATTCTTCAAGCTCAGATACATGTTGGCAATGAACATGACGTTCTG 720
DB 734 GACCCATCAATGAAGATTCTTCAAGCTCAGATACATGTTGGCAATGAACATGACGTTCTG 793
QY 721 CTGCTTGAATGACCTCTCTGAAATGAAACAAATGCACTGACCTTTCAGAAATCTAATGGA 780
DB 794 CTGCTTGAATGACCTCTCTGAAATGAAACAAATGCACTGACCTTTCAGAAATCTAATGGA 853
QY 781 TGGACCAATCCAAAGGCGAATGACCTCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAG 840
DB 854 TGGACCAATCCAAAGGCGAATGACCTCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAG 913
QY 841 ATAGAGAGAAATTATGAATGAAACAAATATTGAGAGCCCTAGGCTGAAAGATATCTTT 900
DB 914 ATAGAGAGAAATTATGAATGAAACAAATATTGAGAGCCCTAGGCTGAAAGATATCTTT 973
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGG 960
DB 974 GATGAATCCAAAGCAGATCTCTCTGGGATTGCTTCGGGGGGTCTCTGTATATATCAAGG 1033
```

```

QY 961 ATGATGCAAAATCTTACATAGAGGTCTCTGAGAGGGGACCGAGGCTACTGCTGCCACA 1020
Db 1034 ATGATGCAAAATCTTACATAGAGGTCTCTGAGAGGGGACCGAGGCTACTGCTGCCACA 1093
QY 1021 GGAAGTAATATGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGGTGACACACCA 1080
Db 1094 GGAAGTAATATGTAGAAAAGCACTCCCTCAGTCCACGCTGTTTAGAGGTGACACACCA 1153
QY 1081 TTCTTATTTGTTATCAGAAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1140
Db 1154 TTCTTATTTGTTATCAGAAAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGCCCT 1213
QY 1141 TGA 1143
Db 1214 TGA 1216

RESULT 5
LOCUS D88575 1950 bp mRNA linear PRI 04-JUL-1998
DEFINITION Homo sapiens mRNA for TP55, complete cds.
ACCESSION D88575
VERSION D88575.1 GI:3288674
KEYWORDS TP55.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Tsujimoto, M., Teurukua, N., Ishida, N., Kurihara, T., Iwasa, F.,
Yamashiro, K., Rogi, T., Kodama, S., Katsuragi, N., Adachi, M.,
Katayama, T., Nakao, M., Yamauchi, K., Hashino, J., Hanyama, M.,
Mura, K., Nakanishi, I., Nakazato, H., Teramura, M., Mizoguchi, H. and
Yamaguchi, N.
TITLE Purification, cDNA cloning, and characterization of a new serpin
with megakaryocyte maturation activity
J. Biol. Chem. 272 (24), 15373-15380 (1997)
97326116
PUBMED 9182567
REFERENCE 2 (bases 1 to 1950)
AUTHORS Tsuruoka, N.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-1996) Nobuo Tsuruoka, Suntory Institute for
Biomedical Research, Pharmaceutical Research Laboratories I; 1-1-1
Wakayamadai, Shimamoto-cho, Mishima-gun, Osaka 618, Japan
(E-mail:tsuruoka@minase.suntory.co.jp, Tel:075-962-9285 (ex.3110),
Fax:075-962-6448)
FEATURES
source
location/Qualifiers
1..1950
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="A431"
74..1216
/cell_type="epidermoid carcinoma"
/codon_start=1
/product="TP55"
/protein_id="BAA31232.1"
/db_xref="GI:3288675"
CDS
translation="MASLAANAAPFCNLFREMDNQGNGVFFSLIFLAALVRL
GAQDLSQDKLLHVTASGNSQSGSQSLKRVFSDINASHKDYDLSVNGL
FAKVYGFHDYIBCAEKLDAKVERVDNLETRNINKVNEVTHGKKNVIG
GGISSAWVLNNAVYKQWQSPKSTINCHPKSPKSCXAVAMHQERKNLSV
IEDSKMLERNGVINYVLLPNDLSIENKLTFCNLNMEWNPFRNRSKYVEVFF
PQFLKXVEMQYLALGLDIPDSKADLSIASGRLYIRNMHKSYLEVEEQT
RATAATGSNIVEKQLFQSTLFRADHPFLFVIRKDIILFSGKVSCT"
BASE COUNT 570 a 384 c 407 g 589 t
ORIGIN
Query Match 100.0%; Score 1143; DB 9; Length 1950;
Best Local Similarity 100.0%; Pred. No. 1e-311;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATGCGCTCCCTTGTCTGAGCAAAATGACAGATTTTGGCTTCAACTGTTTACAGAGATGGAT 60
Db 74 ATGCGCTCCCTTGTCTGAGCAAAATGACAGATTTTGGCTTCAACTGTTTACAGAGATGGAT 133
QY 61 GACAAATCAGGAAATGGAATGTGTTCTTTCCTCTCTGAGAGCTCTTGGCTGCCCTGGCC 120
Db 134 GACAAATCAGGAAATGGAATGTGTTCTTTCCTCTCTGAGAGCTCTTGGCTGCCCTGGCC 193
QY 121 CTGCTCGGCTTGGCGCTCAAGATGACTCCCTCTCTCAGATTTGATAAGTTCTTCATGTT 180
Db 194 CTGCTCGGCTTGGCGCTCAAGATGACTCCCTCTCTCAGATTTGATAAGTTCTTCATGTT 253
QY 181 AACACTGCCCTCAGGATATGGAATCTCTTCTATAGTCACTCAGGCTCAGCTTCAACTG 240
Db 254 AACACTGCCCTCAGGATATGGAATCTCTTCTATAGTCACTCAGGCTCAGCTTCAACTG 313
QY 241 AAMAGAGTTTCTTCTGATATAAATGCATCCACAAAGGATTTATGATCTCAGCATTTGTAAT 300
Db 314 AAMAGAGTTTCTTCTGATATAAATGCATCCACAAAGGATTTATGATCTCAGCATTTGTAAT 373
QY 301 GGGCTTTTCTGAAAAAGTGTATGCTTTTCAATAGGACTACATTTGAGTGTCCGAAAAA 360
Db 374 GGGCTTTTCTGAAAAAGTGTATGCTTTTCAATAGGACTACATTTGAGTGTCCGAAAAA 433
QY 361 TTATACATGCCCAAGTGGAGCGAGTTGACTTTAGCAATCATTTAGAGACACTAGACGT 420
Db 434 TTATACATGCCCAAGTGGAGCGAGTTGACTTTAGCAATCATTTAGAGACACTAGACGT 493
QY 421 AATATTAAATAGTGGTTGAAAATGAACACATGGCAAAATCAAGAACGTGATTTGGTGA 480
Db 494 AATATTAAATAGTGGTTGAAAATGAACACATGGCAAAATCAAGAACGTGATTTGGTGA 553
QY 481 GGTGGCATTAAGCTCATCTGCTGTAATGGTGGTGAATGCTGTGTACTTCAAGGCAAG 540
Db 554 GGTGGCATTAAGCTCATCTGCTGTAATGGTGGTGAATGCTGTGTACTTCAAGGCAAG 613
QY 541 TGGCAATCAGCTTCCACAGAGCGAAACCAATTAATGGCAATTTCAATCTCCCAAGTGC 600
Db 614 TGGCAATCAGCTTCCACAGAGCGAAACCAATTAATGGCAATTTCAATCTCCCAAGTGC 673
QY 601 TCTGGGAGGAGCTCGCCATGATGATCAGGAACGGAAGTTCAATTTCTGTTATTGAG 660
Db 674 TCTGGGAGGAGCTCGCCATGATGATCAGGAACGGAAGTTCAATTTCTGTTATTGAG 733
QY 661 GACCCATCAATCAGATCTCTGAGCTCAGATACAAATGGTGGCATAAAATCAGTACGTTCTG 720
Db 734 GACCCATCAATCAGATCTCTGAGCTCAGATACAAATGGTGGCATAAAATCAGTACGTTCTG 793
QY 721 CTGCTCGAATGAGCTCTCTGGAATGGAACAAACAACTGACCTTTCAGATCTTAATGGA 780
Db 794 CTGCTCGAATGAGCTCTCTGGAATGGAACAAACAACTGACCTTTCAGATCTTAATGGA 853
QY 781 TGGACCAATCCCAAGGCAATGACCTCTAAGTATGTTGAGTATTTTTTTCCTCAGTTCAAG 840
Db 854 TGGACCAATCCCAAGGCAATGACCTCTAAGTATGTTGAGTATTTTTTTCCTCAGTTCAAG 913
QY 841 ATAGAGAGAAATTAATGAATATGAATAAATATTTAGAGCCCTTAGGGCTGAAGATATCTTT 900
Db 914 ATAGAGAGAAATTAATGAATATGAATAAATATTTAGAGCCCTTAGGGCTGAAGATATCTTT 973
QY 901 GATGAATCCCAAGCAGATCTCTCGGATTCCTTTCGGGGGCTCGTCTGTATATCAAGG 960
Db 974 GATGAATCCCAAGCAGATCTCTCGGATTCCTTTCGGGGGCTCGTCTGTATATCAAGG 1033
QY 961 ATGATGCAAAATCTTACATAGAGGTCTCTGAGAGGGGACCGAGGCTACTGCTGCCACA 1020
Db 1034 ATGATGCAAAATCTTACATAGAGGTCTCTGAGAGGGGACCGAGGCTACTGCTGCCACA 1093
QY 1021 GGAAGTAATATGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGGTGACACACCA 1080
Db 1094 GGAAGTAATATGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGGTGACACACCA 1153

```

QY	1081	TTCCATTATTTCTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGGCCT	1140
Db	1154	TTCCATTATTTCTATCAGGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTGGCCT	1213
QY	1141	TGA 1143	
Db	1214	TGA 1216	
RESULT 5			
AF027866			
LOCUS	AF027866	2249 bp mRNA linear PRI 20-OCT-1998	
DEFINITION	Homo sapiens meglin mRNA, complete cds.		
ACCESSION	AF027866		
VERSION	AF027866.1	GI:3769372	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	Miyata, T., Nangaku, M., Suzuki, D., Inagi, R., Uragami, K., Sakai, H.,		
	Okubo, K. and Kurokawa, K.		
TITLE	A meglin-predominant gene, meglin, is a new serpin up-regulated		
	in IgA nephropathy		
JOURNAL	J. Clin. Invest. 102 (4), 828-836 (1998)		
MEDLINE	98376492		
PUBMED	9710452		
REFERENCE	2 (bases 1 to 2249)		
AUTHORS	Toshio, M., Yasuda, Y., Nangaku, M. and Kurokawa, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-OCT-1997) Medical Science, Tokai University,		
	Bousaidai, Isehara, Kanagawa 259-11, Japan		
REFERENCE	3 (bases 1 to 2249)		
AUTHORS	Toshio, M., Yasuda, Y., Nangaku, M. and Kurokawa, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-OCT-1998) Medical Science, Tokai University,		
	Bousaidai, Isehara, Kanagawa 259-11, Japan		
REMARK	Sequence update by submitter		
COMMENT	On Oct 20, 1998 this sequence version replaced gi:3764058.		
FEATURES	Location/Qualifiers		
source	1..2249		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/cell_type="mesangial"		
	365..1507		
	/codon_start=1		
	/product="meglin"		
	/protein_id="AAC64506.1"		
	/db_xref="GI:3769373"		
	/translation="MASLAANAEEFCNLFREMDNQNGNVFSSLSLFAALALVRL		
	GAQDLSQIDKLLHNTASGYGNSNSQSGLSQKRVFSDINASHKDYDLSLVNGL		
	FAEYGFHDKVIECAEKLADKVERVDFNLEEDRETNKWNENHGIKNVIGL		
	GDISSAVMLVNAVYFKWQSAFTKSETINCHFKPKSGKAVAMHQRKFNLSV		
	IEDSKMLLELYNGVINYVLLPNDLSEINIKLTFQNLMEWTNPRMTSKYVEVFF		
	POFXIKNVEMKQYLKGLKIDFDSKDLGLASGRLYISRMHKSYLEVTEBET		
	EATATSGNVKQLPQSLFRADHPFLVIRKODILLPSGKVSCH"		
BASE COUNT	649 a 452 c 461 g 683 t 4 others		
ORIGIN			
Query Match	100.0%; Score 1143; DB 9; Length 2249;		
Best Local Similarity	100.0%; Pred. No. 1.1e-311;		
Matches 1143; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGGCTCCCTGCTGAGCAATGAGATTTGGTTTCACTGTTTCAAGAGATGAT	60
Db	365	ATGGCTCCCTGCTGAGCAATGAGATTTGGTTTCACTGTTTCAAGAGATGAT	424
QY	61	GACATCAAGAAATGAAATGTGTCTTTTCTCTCTGAGCCTCTGCTGCTGCTG	120
Db	425	GACATCAAGAAATGAAATGTGTCTTTTCTCTCTGAGCCTCTGCTGCTGCTG	484
QY	121	CTGGTCCGCTTGGGCTCTCAGATGACTCCCTCTCTCAGATTGATAGTTGCTT	180
Db	485	CTGGTCCGCTTGGGCTCTCAGATGACTCCCTCTCTCAGATTGATAGTTGCTT	544
QY	181	AACACTGCTCAGGATATGGAATCTCTTCTAAATAGTCACTCAGGCTCCAGT	240
Db	545	AACACTGCTCAGGATATGGAATCTCTTCTAAATAGTCACTCAGGCTCCAGT	604
QY	241	AAAAAGATTTTCTGATATTAATGATGCCCAACAGGATTTATGATCTCAGCA	300
Db	605	AAAAAGATTTTCTGATATTAATGATGCCCAACAGGATTTATGATCTCAGCA	664
QY	301	GGGCTTTTGTGCTGAAAAGTGTATGGCTTTTCAATAGGACTACATTGAGTG	360
Db	665	GGGCTTTTGTGCTGAAAAGTGTATGGCTTTTCAATAGGACTACATTGAGTG	724
QY	361	TTATACGATGCCAAAGTGGAGCGAGTTGACTTTTCAATTTAGAACACTAGAG	420
Db	725	TTATACGATGCCAAAGTGGAGCGAGTTGACTTTTCAATTTAGAACACTAGAG	784
QY	421	ATATTTAATAGTGGCTTGAATTAAGAACACATGGCAAAATCAAGACGTGAT	480
Db	785	ATATTTAATAGTGGCTTGAATTAAGAACACATGGCAAAATCAAGACGTGAT	844
QY	481	GTGGGCTAAGCTCATCTCTGTAATGGTGTGCTGTAATGGTGTGTAATGGT	540
Db	845	GTGGGCTAAGCTCATCTCTGTAATGGTGTGTAATGGTGTGTAATGGT	904
QY	541	TGCAATCAGCTTCAACAGAGGCAAAACATATAATTTGCCATTTCAATCTCC	600
Db	905	TGCAATCAGCTTCAACAGAGGCAAAACATATAATTTGCCATTTCAATCTCC	964
QY	601	TCTGGAGAGGAGTCCGCTGATGATCAGACGAGGAGTTCAATTTGTCTGTAT	660
Db	965	TCTGGAGAGGAGTCCGCTGATGATCAGACGAGGAGTTCAATTTGTCTGTAT	1024
QY	661	GACCATCAATGAAGATTCTTGAGCTCAGATACAAATGGTGGCAATAACATG	720
Db	1025	GACCATCAATGAAGATTCTTGAGCTCAGATACAAATGGTGGCAATAACATG	1084
QY	721	CTGCTTGAGATGACCTCTCTGAAATTTGAAAACAAATCTGACCTTTCAGAT	780
Db	1085	CTGCTTGAGATGACCTCTCTGAAATTTGAAAACAAATCTGACCTTTCAGAT	1144
QY	781	TGCAACATCCAGGCGATGACCTCTAGATATGTTTGGGTATTTTCTCTCAGT	840
Db	1145	TGCAACATCCAGGCGATGACCTCTAGATATGTTTGGGTATTTTCTCTCAGT	1204
QY	841	ATAGAAGAATTTATGAATGAAAACAATATTGAGAGCCCTAGAGGCTGAAAG	900
Db	1205	ATAGAAGAATTTATGAATGAAAACAATATTGAGAGCCCTAGAGGCTGAAAG	1264
QY	901	GATGAATCCAAAGCAGATCTCTCGGATTCCTTCGGGGGCTCTGTATATAT	960
Db	1265	GATGAATCCAAAGCAGATCTCTCGGATTCCTTCGGGGGCTCTGTATATAT	1324
QY	961	ATGATGCAAAATCTTACATAGAGTCTCAGGAGGCAACCGAGGCTACTGTG	1020
Db	1325	ATGATGCAAAATCTTACATAGAGTCTCAGGAGGCAACCGAGGCTACTGTG	1384
QY	1021	GGAAGTAATATTGTAGAAAACCAATCCCTCAGTCCACGCTTTTATAGAGT	1080
Db	1385	GGAAGTAATATTGTAGAAAACCAATCCCTCAGTCCACGCTTTTATAGAGT	1444
QY	1081	TTCTTATTTGTTATCAGGAAGATGACATCATCTTATTAGTGGCAAGTTTCT	1140
Db	1445	TTCTTATTTGTTATCAGGAAGATGACATCATCTTATTAGTGGCAAGTTTCT	1504
QY	1141	TGA 1143	
Db	1505	TGA 1507	

## RESULT 7

AF105328 1386 bp mRNA linear ROD 19-OCT-2001  
 Mus musculus mgsin mRNA, complete cds.  
 DEFINITION  
 AF105328  
 VERSION  
 AF105328.1 GI:16269518  
 KEYWORDS  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1386)  
 Nangaku, M., Miyata, T., Suzuki, D., Umezono, T., Hashimoto, T.,  
 Wada, T., Yagi, M., Nagano, N., Inagi, R. and Kurokawa, K.  
 Cloning of rodent mgsin revealed its up-regulation in  
 mesangio proliferative nephritis  
 JOURNAL  
 MEDLINE  
 21368006  
 PUBMED  
 11473647  
 REFERENCE  
 2 (bases 1 to 1386)  
 Miyata, T., Inagi, R. and Kurokawa, K.  
 Direct Submission  
 TITLE  
 Submitted (09-NOV-1999) Institute of Medical Sciences, Tokai  
 University, Bohseidai, Isehara, Kanagawa 259-1193, Japan  
 JOURNAL  
 Location/Qualifiers  
 FEATURES  
 source  
 1..1386  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /cell\_type="mesangial"  
 1..1143  
 /note="human mgsin homolog"  
 /codon\_start=1  
 /product="mgsin"  
 /protein\_id="AA146768.1"  
 /db\_xref="GI:16269519"  
 /translation="MASLAANAEEFGDLFREMDSOGNGVFFSLSIPTALILRL  
 GARDCAROIKALHFNIPSSGNSNPGLOYOLXRVLADINSHKDYELSLATGV  
 FAEKVDFHDIIECAENLNAKVERBFDNDVDFKINKWENHMKVILGSD  
 SSUSGAVNLVNAVIFNGKSAFTKDTLSRCFRSFTCPGVNVMVHQRRENL  
 IQPPQVLELYHGGLSWINEPEDGLCEIESKLSQNLMTNRRKQSVYVNF  
 POKPKENTEMHLKSLGLKIDFESSADLGSAGGRLYSKLMHKSFIENSEGT  
 EATAANNIVEKOLPESTVFRAADRPFLVFKNDIILFTGVKSCP"

## CDS

BASE COUNT 407 a 293 c 296 g 393 t 1 others  
 ORIGIN  
 Query Match 68.4%; Score 781.4; DB 10; Length 1386;  
 Best Local Similarity 80.2%; Pred. No. 1.7e-209;  
 Matches 917; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTCTGCTCAGCAATGCGAGCTTTTGTCTCAACCTGTTCTCAGAGATGGAT 60  
 DB |||||  
 QY 1 ATGGCTCCCTTCTGCTCAGCAATGCGAGCTTTTGTCTTGGACTTTTTCAGAGATGGAT 60  
 DB |||||  
 QY 61 GACAATCAAGAAATGGAATGTGTTCTTTCTCTCTGAGCTCTGCTGCGCTGACC 120  
 DB |||||  
 QY 61 ACTAGCCAGGAATGGAATGTATTTCTCTCTCTGAGCTCTGCTGCGCTGACC 120  
 DB |||||  
 QY 121 CTGGTCCGCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATGATGAAGTTGCTTCATGTT 180  
 DB |||||  
 QY 121 CTAATCCGCTGGGTGCTCGAGTGACTGTGACGTCAGATTGACAGGCACTGCACTTT 180  
 DB |||||  
 QY 181 AACATGCTCCAGGATATGGAATCTTCTAATAGTCAGTCAGGCTCAGCTCAACTG 240  
 DB |||||  
 QY 181 AACATGCAATCAAGACAGGAACTCATCTAATATATCAGCCAGGACTTCAGTATCAATTG 240  
 DB |||||  
 QY 241 AAAGAGTTTTTCTGATATAAATGCAATCCCAAGATATATGATCTCAGCATTTGAT 300  
 DB |||||  
 QY 241 AAAGAGTTTTTCTGATATAAATGCAATCCCAAGATATATGATCTCAGCATTTGAT 300  
 DB |||||  
 QY 301 GGGCTTTTGTGAAAAGTGTATGGCTTTCATAGGACTACATGAGTGTCCGAAAAA 360  
 DB |||||  
 QY 301 GGAGTTTTTGGAGAAAAAGTCTATGACTTTTCATAGAACTACATGAGTGTCTGAAAC 360  
 DB |||||

QY 361 TTATCAGATGCCAAGTGGAGCGAGTGTGACTTTTACGATCATTTTGAAGACACTAGACCT 420  
 DB |||||  
 QY 361 TTATCAATGCTAAAGTGGAAAGAGTGTGACTTTCACAAATGATGTACAAGATACCAAGATT 420  
 DB |||||  
 QY 421 AATATTAAATAGTGGTTGAAATGAAACACATCGGCAAAATCAAGACGTCATTGGTGA 480  
 DB |||||  
 QY 421 AATATTAAATAGTGGTTGAAATGAAACACATCGGCAAAATCAAGACGTCATTGGTGA 480  
 DB |||||  
 QY 481 GGTGGCATTAAGTCATCTGCTGTAATGGTGTGTAATGGTGTGTAATGGTGTGTAATGG 540  
 DB |||||  
 QY 481 AGCAGGCTCAGCTCGTGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 DB |||||  
 QY 541 TGGCATTAAGCTTTCACCAAGAGGAAACCAAAATTTGCAATTTCAAAATTTCCCAAGTGC 600  
 DB |||||  
 QY 541 TGGAAATCGGCTTTCACCAAGACTGATACCTCAGTTGGCTTTTAGGTCTCCCAAGTGT 600  
 DB |||||  
 QY 601 TCTGGGAAGGAGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 DB |||||  
 QY 601 CTTGGAAAGTAGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 DB |||||  
 QY 661 GACCATCAATCAAGATTTCTTGGCTCAGATACAAATGGTGGCATAAATGATGATGATGAT 720  
 DB |||||  
 QY 661 CACCCACCAATGCGAGTTCTTGGCTCCTCAATATCATGCTGGCTGATGATGATGATGAT 720  
 DB |||||  
 QY 721 CTCCTGAGATGACCTTCTGAAATTTGAAATGAAATGAAATGAAATGAAATGAAATGAA 780  
 DB |||||  
 QY 721 CTCCTGAGATGACCTTCTGAAATTTGAAATGAAATGAAATGAAATGAAATGAAATGAA 780  
 DB |||||  
 QY 781 TGGACCATCCCAAGGCGATGACCTTAAAGTATGTTGAGGTATTTTCTCAGTTCAAG 840  
 DB |||||  
 QY 781 TGGACCATGAGGAGGAAATGAAATCTCAGTATGTAACGCTTTCTCCCTCCAGTTCAAG 840  
 DB |||||  
 QY 841 ATAGAGAAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 900  
 DB |||||  
 QY 841 ATAGAGAAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 900  
 DB |||||  
 QY 901 GATGAATCAAGCAGATCTCTCTGGGATGCTTTCGGGGGGTCTGCTGATATATCAAGG 960  
 DB |||||  
 QY 901 GATGATCCAGTGCAGATCTCTCTGGGATGCTTTCGGGGGGTCTGCTGATATATCAAGG 960  
 DB |||||  
 QY 961 ATGATGCAAAATCTTACATAGAGTCTACGAGGAGGACCGAGGCTACTGCTGCTGCA 1020  
 DB |||||  
 QY 961 CTATGCAAAATCTTACATAGAGTCTACGAGGAGGACCGAGGCTACTGCTGCTGCA 1020  
 DB |||||  
 QY 1021 GGAAGTAAATATGTAAGAAAGCACTCCCTCAGTCCACGCTTTTAGAGTGCACCAACCA 1080  
 DB |||||  
 QY 1021 GGAATAAATATGTAAGAAAGCACTCCCTCAGTCCACGCTTTTAGAGTGCACCAACCA 1080  
 DB |||||  
 QY 1081 TTCTTATTTTATCAGGAGGATGATCATATCTTATTCAGTGGCAAGTTTCTTGGCT 1140  
 DB |||||  
 QY 1081 TTCTTATTTTATCAGGAGGATGATCATATCTTATTCAGTGGCAAGTTTCTTGGCT 1140  
 DB |||||  
 QY 1141 TGA 1143  
 DB |||||  
 QY 1141 TGA 1143  
 DB |||||

## RESULT 8

AF105329 1143 bp mRNA linear ROD 19-OCT-2001  
 LOCUS  
 DEFINITION  
 Rattus norvegicus mgsin mRNA, complete cds.  
 AF105329  
 ACCESSION  
 AF105329.1 GI:16269520  
 VERSION  
 KEYWORDS  
 SOURCE  
 Rattus norvegicus (Norway rat)  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 1143)  
 Nangaku, M., Miyata, T., Umezono, T., Hashimoto, T.,  
 Wada, T., Yagi, M., Nagano, N., Inagi, R. and Kurokawa, K.



/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
BASE COUNT 378 a 243 c 276 g 332 t  
ORIGIN

Query Match 66.1%; Score 755.8; DB 6; Length 1229;  
Best Local Similarity 78.8%; Pred. No. 2.9e-202;  
Matches 901; Conservative 0; Mismatches 242; Indels 0; Gaps 0;  
QY 1 ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 60  
Db ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 67  
QY 61 GCAATCAAGGAATGGAATGTTCTTCTCTCTGAGCTCTTCGCTCCCTGACC 120  
Db AGTAGTCAAGGAATGGAATGTTCTTCTCTCTGAGCTCTTCGCTCCCTGACC 127  
QY 121 CTGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 180  
Db ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 187  
QY 128 CTGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 240  
Db ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 247  
QY 241 AAGAGATGTTTTCATATATAAATGATCCACAGGATTAATGATCTGAGATGAT 300  
Db AAGAGATGTTTTCATATATAAATGATCCACAGGATTAATGATCTGAGATGAT 307  
QY 301 GGGCTTTTCTGAAAGATGATGGCTTTTCAATGAGTACATGAGTGTCCGAAAAA 360  
Db GGGCTTTTCTGAAAGATGATGGCTTTTCAATGAGTACATGAGTGTCCGAAAAA 367  
QY 361 TTATAGATGCAAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 420  
Db TTATAGATGCAAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 427  
QY 421 AATATTAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 480  
Db AATATTAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 487  
QY 481 GGTGATCAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 540  
Db GGTGATCAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 547  
QY 541 TGGCAATCAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 600  
Db TGGCAATCAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 607  
QY 601 TCTGGAGGATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 660  
Db TCTGGAGGATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 667  
QY 661 GACCATCAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 720  
Db GACCATCAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 727  
QY 721 CTGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 780  
Db ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 787  
QY 781 TGGCAATCAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 840  
Db TGGCAATCAATGAGTGTGAGGATGTTTCAATGAGTACATGAGTGTCCGAAAAA 847  
QY 841 ATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db ATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907  
QY 901 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967

QY 961 ATGATGCAAAATCTTACATAGAGTCTACAGAGGACCGAGGCTACTGCTGCCACA 1020  
Db CTAAATGCAAAATCTTACATAGAGTCTACAGAGGACCGAGGCTACTGCTGCCACA 1027  
QY 1021 GGAATATATTTAGAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGTGCACCA 1080  
Db GGAATATATTTAGAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGTGCACCA 1087  
QY 1081 TTCTATTTTATTCAGAGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTCT 1140  
Db TTCTATTTTATTCAGAGAGGATGACATCATCTTATTCAGTGGCAAGTTCTTCT 1147  
QY 1141 TGA 1143  
Db 1148 TGA 1150  
RESULT : 0  
LOCUS BD094241 1229 bp DNA linear PAR 27-AUG-2002  
DEFINITION Model animals of nephropathy proliferating mesangial cells.  
ACCESSION BD094241  
VERSION BD094241.1 GI:22639829  
KEYWORDS WO 0124628-A/2.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1229)  
AUTHORS Miyata, T.  
TITLE Model animals of nephropathy proliferating mesangial cells  
JOURNAL Patent: WO 0124628-A 2 12-APR-2001;  
TOSHIO MIYATA, KIYOSHI KUROKAWA  
COMMENT OS Rattus norvegicus (rat)  
PN WO 0124628-A/2  
PD 12-APR-2001  
PF 06-OCT-2000 WO 2000JP006988  
PR 06-OCT-1999 JP 99P 285736  
PI TOSHIO MIYATA  
PC A01K67/027, A61P13/12, A61K45/00, C12N15/12, C12N15/85, G01N33/15,  
G01N33/50  
CC  
FH  
FT  
Key CDS Location/Qualifiers  
1..1229  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
BASE COUNT 377 a 242 c 274 g 330 t 6 others  
ORIGIN  
Query Match 65.9%; Score 753; DB 6; Length 1229;  
Best Local Similarity 78.5%; Pred. No. 1.8e-201;  
Matches 897; Conservative 0; Mismatches 246; Indels 0; Gaps 0;  
QY 1 ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 60  
Db ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 67  
QY 61 GCAATCAAGGAATGGAATGTTCTTCTCTCTGAGCTCTTCGCTCCCTGACC 120  
Db AGTAGTCAAGGAATGGAATGTTCTTCTCTCTGAGCTCTTCGCTCCCTGACC 127  
QY 121 CTGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 180  
Db ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 187  
QY 128 CTGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 240  
Db ATGGCTCCCTCTCCAGCAATGCAAGTCTTTCCTCAACCTGTTCAGAGATGAT 247  
QY 181 AACCTGCTCCAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
Db AACCTGCTCCAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 247

	SOURCE	Mus musculus (house mouse)
	ORGANISM	Mus musculus
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Rutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1147)
	AUTHORS	Miyata,T.
	TITLE	Model animals of nephropathy proliferating mesangial cells
	JOURNAL	Patient: WO 0124628-A 3 12-APR-2001;
	COMMENT	TOSHIO MIYATA,KIYOSHI KUROKAWA OS Mus musculus [mouse] PN WO 0124628-A/3 PD 12-APR-2001 PF 06-OCT-2000 WO 2000JP006988 PR 06-OCT-1999 JP 99P 285736 PT TOSHIO MIYATA PC AUKES7/027,A6IP13/12,A6IK45/00,C12N15/12,C12N15/85,G0IN33/15, G0IN33/50 CC CC FH Key Location/Qualifiers FT CDS (1)..(1104).
	FEATURES	source 1..1147 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090"
	BASE COUNT	351 a 239 c 250 g 307 t
	CSIGIN	
	Query Match	65.5%; Score 748.6; DB 6; Length 1147;
	Best Local Similarity	79.8%; Pred.No.3.2e-200;
	Matches	883; Conservative 0; Mismatches 224; Indels 0; Gaps 0;
Qy	241	AAAAGAGTTTTTCTGATATAATGCATCCCAAGGATTATGATCTCAGCATTTGGAAAT 300
Dd	248	AAAAGAGTTCTTGTCACATAAACCTCATCTCAPAAGGATNNNAATCTCAGCATTTGCCAAT 307
Qy	301	GSGCTTTTGTCTGAAGAAGTAGTGCTTTTCATPAAGGACTACATTGAGTGTGCGRAAAA 350
Dd	308	GGAGTTTTTGAGGAAGAAGTAATTGATTTTCATAGAGCTATATGAGGTGTCTGGAAC 357
Qy	361	TTATACGATGCGAAAGTGGAGCGAGTTGACATTTTACGAATCATTTTAGAGACAATAGACGT 420
Dd	368	TTATACANTGCTAAAGTGGAAAGAGTTGATTTTACAAATGATATACAGAACAACAGATTIT 427
Qy	421	AATATTAATAGTGGTTGMAATGAACACATGCGAAAAATCAAGAACCTGATTTGTTCAR 480
Dd	428	APAATTAATAAATGATTTGAAAATGAACACATGCGCAAATCAAGAAGGTGTGTGGGGAC 487
Qy	481	GSTGGCATAAAGCTCATCTGCTGAATGTGCTGGTGAATGTGTGATCTTCAAAGGCAAG 540
Dd	488	ASCAGCCTCACATCNAGTCTCANSGTCTAGTGTAATGCTGTTTACTTCAAAGCCAAG 547
Qy	541	TGCAATACAGCTTACCAGAGCGGAACCAATTAATTCGCAATTCAMAATCTCCMAATGTC 600
Dd	548	TGGAATCGGCTTACCAGAGTAGTATACCTCAGTGTGCATTTCAAGTCTCCACGCGGT 607
Qy	601	TCTGGGAGGCGAGTGGCCATGATCATCAGAAAGGAAGTTCAATTTGTCTGTATTAG 660
Dd	608	CCTGMAAGANGTAAATNTGATGCTNCAGAACGAGGTTCAMTTGTCTACCAATTCAG 667
Qy	661	GACCCATCAATGAAGATCTTGAGCTCAGATACAATPGTGGCATAAACATGPAOTGCTG 720
Dd	668	GAGCCACCAATGCGAGATCTTCAGCTCAATATCATGTGGCATAGCATGTACATCATG 727
Qy	721	CTSCCTGAGATGACCTCTGMAATTGABAAACAACTGACCTTCAGBAATCTAATGSA 750
Dd	728	TTGCCCGAGATGACCTATCCGAAATGAAAGCAAGCTGAGTTTCCAGAAATCTAATGGAC 757
Qy	781	TGACCAATCCOAGGCGAATGACCTCTAAGTATGTTGAGGTATTTTTTCTCAGTTCAG 840
Dd	788	TGACCAATAGCAGGAAGATGAAATCTCAGTATGTGAATGTGTTCTCCCGCAGTTCAAG 847
Qy	841	ATAGACAGAAATTTGAATGAAACAAATTTTGAAGGCCCTTAGGGCTGAAAGATCTTT 900
Dd	848	ATAGAGAAAGATTTAGAAATGAGGAGCCACTTGAATCTGTAGGCTTGGAGACATCTTT 907
Qy	901	GATGAATCCAAAGCAGATCTCTCTGGGATTCCTTGGGGGGTGCTGTGTATATATCAAG 960
Dd	908	GTTGAGTCCAGGGTGATCTGTCTGGAATTCGCTCTGGAGGTGCTCTATGATCAAG 967
Qy	961	ATGATGCACAAATCTTACATAGAGTGTCTGAGAGGGGACCGAGGCTACTGTCGCACA 1020
Dd	968	CTAATCCAAAGTCCCTCATFAGAGTCTCAGAAGAGGCCGAGGCAACTGCTGCCACA 1027
Qy	1021	GGAAGTAATTTGAGAAAGAACTCCCTCAGTCCACCGCTGTTAGAGCTGACCAACCA 1080
Dd	1028	GAAAGTAACTCTGTBAAAGCTACTTCTGAAATCCAAGGCTTCAGAGCTGACCGCCC 1087
Qy	1081	TTCTATTTTGTATCAGGAGGAGCAATCATTTTTCAGTGGCAAGTTTCTTGCCCT 1140
Dd	1088	TTTCTGTTGTATTAGGAAGATGGCATCTTTATTTACTTGGCAAGTCTCGTGTCT 1147
Qy	1141	TGA 1143
Dd	1148	TGA 1150
	RESULT 11	
	BD094242	
	LOCUS	1147 bp DNA linear PAT 27-AUG-2002
	DEFINITION	Model animals of nephropathy proliferating mesangial cells.
	ACCESSION	BD094242
	VERSION	BD094242.1 GI:22639830
	KEYWORDS	WO 0124628-A/3.





```

/clone="RP11-79D21"
/clone_lib="RPC1-11 Human Male BAC"
599. .635
/rpt_family=" (T)n"
repeat_region
repeat_region
1040. .1077
/rpt_family="AT_rich"
repeat_region
1411. .1747
/rpt_family="L1P"
repeat_region
2632. .2992
/rpt_family="MLT1A1"
repeat_region
3115. .3138
/rpt_family="AT_rich"
repeat_region
3487. .3828
/rpt_family="L1MC4a"
repeat_region
3834. .4338
/rpt_family="MLT1C"
repeat_region
4339. .5397
/rpt_family="L1MC4a"
repeat_region
5742. .5899
/rpt_family="MER5B"
repeat_region
6499. .6816
/rpt_family="HAL1"
repeat_region
7416. .7621
/rpt_family="MER3"
repeat_region
8163. .9090
/rpt_family="L2"
repeat_region
9276. .9419
/rpt_family="MIR"
repeat_region
9598. .9658
/rpt_family="MER81"
repeat_region
9659. .10033
/rpt_family="MLT1J"
repeat_region
13986. .14009
/rpt_family="AT_rich"
repeat_region
14011. .14115
/rpt_family=" (TA)n"
repeat_region
14942. .15044
/rpt_family="MIR3"
repeat_region
15156. .15298
/rpt_family="L2"
repeat_region
15592. .15773
/rpt_family="L1M4C"
repeat_region
15974. .18310
/rpt_family="L1PA16"
repeat_region
16373. .16374
/notes="1328 bp of bacterial transposon insertion in BAC
excised between these 2 bp"
18311. .18331
/rpt_family=" (TAAAA)n"
repeat_region
18332. .19566
/rpt_family="L1PA16"
repeat_region
19567. .19587
/rpt_family=" (CA)n"
repeat_region
19588. .20042
/rpt_family="L1PA16"
repeat_region
20043. .20064
/rpt_family="AT_rich"
repeat_region
21259. .21386
/rpt_family="MIR3"
repeat_region
22395. .22420
/rpt_family="AT_rich"
repeat_region
22788. .22948
/rpt_family="FRM"
repeat_region
23843. .24253
/rpt_family="MST"
repeat_region
24678. .25041
/rpt_family="MLT1A2"
repeat_region
25054. .25094
/rpt_family="HAL1"
repeat_region
25214. .25358
/rpt_family="L1M4C"
repeat_region
25602. .25854
/rpt_family="MIR"

```

```

repeat_region 26746. .26796
/rpt_family="AT_rich"
repeat_region 27451. .27484
/rpt_family=" (CAAAA)n"
repeat_region complement(28183. .28219)
/rpt_family="MIR"
repeat_region complement(28249. .28325)
/rpt_family="MIR"
repeat_region complement(29372. .29459)
/rpt_family="MIR"
repeat_region 29539. .30003
/rpt_family="MER66B"
repeat_region 30596. .30624
/rpt_family="AT_rich"
repeat_region 31418. .31829
/rpt_family="L2"
repeat_region 31846. .31915
/rpt_family=" (TGAA)n"
repeat_region complement(32019. .32103)

Query Match 34.9%; Score 399.4; DB 9; Length 157284;
Best Local Similarity 99.8%; Pred. No. 2.8e-101;
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 743 AAATTGAAACAACTGACCTTCAGAACTCTAATGATGACCAATCCAGGCGAATGA 802
Db 126113 AGATTGAAACAACTGACCTTCAGAACTCTAATGATGACCAATCCAGGCGAATGA 126054

QY 803 CCTCTAAGTATGTTGAGGTATTTTCTCAGTTCAGATAGAGAAATTAATGAATGA 862
Db 126053 CCTCTAAGTATGTTGAGGTATTTTCTCAGTTCAGATAGAGAAATTAATGAATGA 125994

QY 863 ALCANATTTGAGAGCCCTAGGCTGAAAGATATCTTTGATGATCAAGCAGATCTCT 922
Db 125993 AACATATTTGAGAGCCCTAGGCTGAAAGATATCTTTGATGATCAAGCAGATCTCT 125934

QY 923 CTGGGATTTGCTCGGGGGTGTCTGTATATATCAAGATGATGACCAATCTTACATAG 982
Db 125933 CTGGGATTTGCTCGGGGGTGTCTGTATATATCAAGATGATGACCAATCTTACATAG 125874

QY 983 AGGTCACCTGAGAGGCGCACCGAGGCTACTGCTGCCACAGGAAGTAATATTGTAGAAAGC 1042
Db 125873 AGGTCACCTGAGAGGCGCACCGAGGCTACTGCTGCCACAGGAAGTAATATTGTAGAAAGC 125844

QY 1043 AACTCCCTAGTCCAGCTGTGTTAGCTGACACCCATTCCTATTGTTGATCAGGAAGG 1102
Db 125813 AACTCCCTAGTCCAGCTGTGTTAGCTGACACCCATTCCTATTGTTGATCAGGAAGG 125754

QY 1103 ATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTGA 1143
Db 125753 ATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCTTGA 125713

RESULT 13
AC027524
LOCUS 189092 bp DNA linear HTS 27-MAR-2003
DEFINITION Homo sapiens chromosome 18 clone RP11-317G1 map 18, 3 unordered
pieces.
ACCESSION AC027524
VERSION AC027524.4 GI:21307437
KEYWORDS HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 189092)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone RP11-317G1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 189092)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,

```

Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPhee, R., Meldrum, J., Meneses, L., Minova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rocov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stancu, S., Stancu, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 189092)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneses, L., Minova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Risse, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tophan, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 1, 2002 this sequence version replaced gi:11990731.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
Project Information  
Center project name: 18516  
Center clone name: 317\_G\_

\* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 85659: contig of 85659 bp in length  
\* 85660 85759: gap of 100 bp  
\* 85760 163014: contig of 77255 bp in length  
\* 163015 163115: gap of 100 bp  
\* 163115 189092: contig of 25978 bp in length.

FEATURES  
source  
Location/Qualifiers  
1. 189092  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18"  
/clone="RP11-317G1"  
/clone\_lib="RP11-11 Human Male BAC"  
BASE COUNT 57107 a 35683 c 36719 g 59369 t 214 others  
ORIGIN

Query Match 34.9%; Score 399.4; DB 2; Length 189092;  
Best Local Similarity 99.8%; Pred. No. 2.8e-101;  
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 743 AAATGAAACAACTGACCTTTTCAGAACTTAATGGAATGACCAATCCAGGCGGAATGA 802  
Db 96044 AGATTGAAACAACTGACCTTTTCAGAACTTAATGGAATGACCAATCCAGGCGGAATGA 96103

QY 803 CCTCTAAGTATCTTGAGGTATTTTTCCTCAGTTCAAGATAGAGAGATTAATGAAATGA 862  
Db 96104 CCTCTAAGTATCTTGAGGTATTTTTCCTCAGTTCAAGATAGAGAGATTAATGAAATGA 96163

QY 863 AACAAATTTTGAGAGCCCTTAGGCTGAAGATATCTTTTGATGAATCCAAAGCAGATCTCT 922  
Db 96164 AACAAATTTTGAGAGCCCTTAGGCTGAAGATATCTTTTGATGAATCCAAAGCAGATCTCT 96223

QY 923 CTGGGATTCCTTCGGGGGGTCTGCTCTATATATCAAGATAGAGAGATTAATGAAATGA 982  
Db 96224 CTGGGATTCCTTCGGGGGGTCTGCTCTATATATCAAGATAGAGAGATTAATGAAATGA 96283

QY 983 AGTCTACTGAGGAGGACCGAGGCTTCTGTCGACAGGAGTAAATTTCTAGAAAGC 1042  
Db 96284 AGTCTACTGAGGAGGACCGAGGCTTCTGTCGACAGGAGTAAATTTCTAGAAAGC 96343

QY 1043 AACTCCCTCAGTCCACGCTGTTTAGAGCTGACCCACCATTCCTATTCTTTTATCAGGAGG 1102  
Db 96344 AACTCCCTCAGTCCACGCTGTTTAGAGCTGACCCACCATTCCTATTCTTTTATCAGGAGG 96403

QY 1103 ATGACATCATCTTATTCAGTGCGCAAGTTTCTTCCCTTGA 1143  
Db 96404 ATGACATCATCTTATTCAGTGCGCAAGTTTCTTCCCTTGA 96444

RESULT 14  
AR270261  
LOCUS AR270261 316 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 824 from patent US 6500938.  
ACCESSION AR270261  
VERSION AR270261.1 GI:29701495  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.  
REFERENCE 1 (bases 1 to 316)  
AUTHORS Au-Young, J. and Seilhamer, J.J.  
TITLE Composition for the detection of signaling pathway gene expression  
JOURNAL Patent: US 6500938-A 824 31-DEC-2002;  
FEATURES  
Location/Qualifiers  
source 1. 316  
/organism="unknown"  
BASE COUNT 92 a 63 c 72 g 88 t 1 others  
ORIGIN

Query Match 26.6%; Score 304; DB 6; Length 316;  
Best Local Similarity 99.4%; Pred. No. 1.4e-74;  
Matches 315; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 822 ATTTTTCCTCAGTCCAGATAGAGAGATTAATGAAATGAAACAAATTTTGTAGAGCCCT 881  
Db 1 ATTTTTCCTCAGTCCAGATAGAGAGATTAATGAAATGAAACAAATTTTGTAGAGCCCT 60

TITLE  
JOURNAL  
COMMENT

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

```

QY 882 AGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCGGAGTTCCTTGGGGGG 941
Db 61 AGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTCTGGGTTGCTTGGGGGG 120
QY 942 TCGTCTGTATATCAAGGATGATGCACAAATCTTACATAGAGTCACTGAGGAGGCAC 1001
Db 121 TCGTCTGTATATCAAGGATGATGCACAAATCTTACATAGAGTCACTGAGGAGGCAC 180
QY 1002 CGAGCTACTGCTGCGACAGGAAGTAATATTGTAGAAAGCAATCCCTCAGTCCACGCT 1061
Db 181 CGAGCTACTGCTGCGACAGGAAGTAATATTGTAGAAAGCAATCCCTCAGTCCACGCT 240
QY 1062 GTTTAGAGCTGACCCACCATCTTATTTGTATCAGGAGGATGACATCATCTATTTCAG 1121
Db 241 GTTTAGAGCTGACCCACCATCTTATTTGTATCAGGAGG-TCACATCATCTATTTCAG 299
QY 1122 TGGCAAAAGTTTCTTGGC 1138
Db 300 TGGCAAAAGTTTCTTGGC 316

RESULT 15
AC125314 121141 bp DNA linear ROD 26-MAR-2003
DEFINITION Mus musculus chromosome 1 clone RP24-361P7, complete sequence.
ACCESSION AC125314
VERSION AC125314.3 GI:29244780
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 121141)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 121141)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Direct Submission
REFERENCE 3 (bases 1 to 121141)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Submitted (22-JUN-2002) Genome Sequencing Center, 4444 Forest Park
JOURNAL Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 121141)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Submitted (03-DEC-2002) Genome Sequencing Center, 4444 Forest Park
JOURNAL Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 121141)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Submitted (26-MAR-2003) Genome Sequencing Center, 4444 Forest Park
JOURNAL Parkway, St. Louis, MO 63108, USA
COMMENT On Mar 26, 2003 this sequence version replaced gi:26006669.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BB0361P07
-----
FEATURES
source
1. 121141
/organism="Mus musculus"
/mol_type="Genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/clone="RP24-361P7"
BASE COUNT 36618 a 23200 c 23524 g 37799 t
ORIGIN

```

Query Match 24.4%; Score 279.4; DS 10; Length 121141;  
 Best Local Similarity 81.0%; Pred. No. 2.3e-57;

```

Matches 325; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 743 AAATTGAAAACAAACTGACCTTTTCAGAAATCTAATGGAATGGACCAATCCAAAGGCGAATGA 802
Db 18914 AGATTGAAAAGCAAGCTGAGTTTCCAGAAATCTGATGAGCTGGACCAATAGGAGGAAAATGA 18973
QY 803 CCTCTAAGTATGTTGAGGTATTTTTCCTCAGTTCAAGATAGAGAAGAAATTTGAATGA 862
Db 18974 AATCTCAGTATGTGAACGTTTCTCCCCCAGTTCAAGATAGAGAAGAAATTTGAATGA 19033
QY 863 AACAAATATTGACAGCCCTTAGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCT 922
Db 19034 CGACCACTTGAATCTTTAGGCTTTGAAGATATCTTTGATGAGTCCAGTGCAGATCTCT 19093
QY 923 CTGGGATTCGTTGGGGGGTGTCTGTATATATATATATATATATATATATATATATATAT 982
Db 19094 CTGGAAATTCCTCTGGAGTCTGTCTCTACGTATCAAAAGCTAATGCACAGTCAATTCATAG 19153
QY 983 AGGTCACTGAGGAGGCGACCGAGGCTACTGTCGCACAGGAGTAATATTGTAGAAAGC 1042
Db 19154 AGGTCTCAGAGGAGGCGACTGAAGCCACTGTCGCACAGAAAATAACATTGTGTGAAGAGC 19213
QY 1043 AACTCCCTCAGTCCACCGCTGTTTAGAGCTGACCAACCCATTCCTATTTTGTTATCAGGAAG 1102
Db 19214 AGCTTCTGAGTCCACAGTGTTCAGAGCGCGACCGCCCTTCTCTGTTGTATCAACAAGA 19273
QY 1103 ATGACATCATCTTATTCAGTGGCAAGTTTCTTGGCCTTGA 1143
Db 19274 ATGACATCATCTTATTTACTGGCAAGTCTCTTGTCTCTGA 19314

Search completed: December 14, 2003, 02:32:56
Job time : 4381 secs

```

**X**

PT Megsin protein expressed specifically in mesangial cells  
 XX Claim 4; Page 58-61; 100pp; Japanese.  
 XX This invention describes the isolation of novel megin nucleic acid and  
 CC proteins from human, rat and mouse tissue. This protein is expressed  
 CC specifically in mesangial cells. The products of the invention are  
 CC useful for the treatment and diagnosis of diseases involving mesangial  
 CC cells, such as IgA nephropathy.  
 XX  
 SQ Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;  
 Query Match 100.08; Score 1143; DB 20; Length 1143;  
 Best Local Similarity 100.08; Pred. No. 0;  
 Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGCTCCCTGCTGACGCAATGCGAGATGTTTGGCTTCAACCTGTTGAGAGATGGAT 60  
 DB 1 ATGGCTCCCTGCTGACGCAATGCGAGATGTTTGGCTTCAACCTGTTGAGAGATGGAT 60  
 QY 61 GACAAATCAAGAAATGGAATGTTGTTTCCCTCTGAGCCTCTTCGCTGCGCTGGCC 120  
 DB 61 GACAAATCAAGAAATGGAATGTTGTTTCCCTCTGAGCCTCTTCGCTGCGCTGGCC 120  
 QY 121 CTGCTCGCTTGGGCGCTCAGATGACTCCCTCTCTCAGATTGATGTTGCTTCATGTT 180  
 DB 121 CTGCTCGCTTGGGCGCTCAGATGACTCCCTCTCTCAGATTGATGTTGCTTCATGTT 180  
 QY 181 AACACTGCTCAGGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240  
 DB 181 AACACTGCTCAGGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240  
 QY 241 AAAAGATTTTTCTGATTAATGATGATCCACAGATTAATGATCTCAGATGTTGCAAT 300  
 DB 241 AAAAGATTTTTCTGATTAATGATGATCCACAGATTAATGATCTCAGATGTTGCAAT 300  
 QY 301 GGGCTTTTCTGMAAAAGTGTATGGCTTCTAAGAGTACATTCAGTGTGCGCAAAA 360  
 DB 301 GGGCTTTTCTGMAAAAGTGTATGGCTTCTAAGAGTACATTCAGTGTGCGCAAAA 360  
 QY 361 TTATACATGCCAAAGTGGAGTGACTTACGAATCAATTTAGAGACACTAGACGT 420  
 DB 361 TTATACATGCCAAAGTGGAGTGACTTACGAATCAATTTAGAGACACTAGACGT 420  
 QY 421 AATATTATTAAGTGGTTGAAATCAACACATGCGCAAAATCAAGACGTGATGTTGAA 480  
 DB 421 AATATTATTAAGTGGTTGAAATCAACACATGCGCAAAATCAAGACGTGATGTTGAA 480  
 QY 481 GGTGGCATAGCTTCATCTGCTGTAATGGTGTGCTGTAATGCTGTACTTCAAGGCAAG 540  
 DB 481 GGTGGCATAGCTTCATCTGCTGTAATGGTGTGCTGTAATGCTGTACTTCAAGGCAAG 540  
 QY 541 TGGCAATCAGCTTCACAGAGGCAAAACCAATTAATGCTCAATCTCCCAAGTGC 600  
 DB 541 TGGCAATCAGCTTCACAGAGGCAAAACCAATTAATGCTCAATCTCCCAAGTGC 600  
 QY 601 TCTGGAGGAGCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 DB 601 TCTGGAGGAGCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 QY 661 GACCCATCAATGAGATCTTGGCTCAGATCAATGCTGGCATAAACATGATGATGATG 720  
 DB 661 GACCCATCAATGAGATCTTGGCTCAGATCAATGCTGGCATAAACATGATGATGATG 720  
 QY 721 CTGGCTGAGATGACCTCTCTGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 780  
 DB 721 CTGGCTGAGATGACCTCTCTGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 780  
 QY 781 TGGACCAATCAAGGCGAATGACCTCTTAAGTATGTTGAGTATTTTTTCTCAGTCAAG 840  
 DB 781 TGGACCAATCAAGGCGAATGACCTCTTAAGTATGTTGAGTATTTTTTCTCAGTCAAG 840  
 QY 841 ATAGAGAGAAATTAATGAAATGAAATTAATTTGAGAGCCCTGAGGCTGAAAGATATCTTT 900

Db 841 ATAGAGAGAAATTAATGAAATGAAATTAATTTGAGAGCCCTGAGGCTGAAAGATATCTTT 900  
 QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTCGGGGGCTCTCTCTGATATATCAAG 960  
 Db 901 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTCGGGGGCTCTCTCTGATATATCAAG 960  
 QY 961 ATGATGACAAATCTTACATAGAGTCACTGAGGAGGCAACGAGGCTACTGCTGCCACA 1020  
 Db 961 ATGATGACAAATCTTACATAGAGTCACTGAGGAGGCAACGAGGCTACTGCTGCCACA 1020  
 QY 1021 GGAGTATATATGTTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTATAGCTGACACCCA 1080  
 Db 1021 GGAGTATATATGTTAGAAAAGCAACTCCCTCAGTCCAGCTGTTTATAGCTGACACCCA 1080  
 QY 1081 TTCTATTTTATATAGGAGGATGATCATCTTATTCAGTGGGAAAGTTCTTGGCCT 1140  
 Db 1081 TTCTATTTTATATAGGAGGATGATCATCTTATTCAGTGGGAAAGTTCTTGGCCT 1140  
 QY 1141 TGA 1143  
 Db 1141 TGA 1143

RESULT 2  
 AAA9294  
 ID AAA9294 standard; cDNA; 1143 BP.  
 XX  
 AC AAA9294;  
 XX  
 DT 30-JAN-2001 (first entry)  
 XX  
 DE Human megin nucleotide sequence SEQ ID NO:1.  
 KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;  
 KW IgA; immunoglobulin A; detection; renal function; renal disorder;  
 KW diagnosis; biological sample; blood; urine; ss.  
 XX Homo sapiens.  
 XX WO200057189-A1.  
 XX 28-SEP-2000.  
 XX 17-MAR-2000; 2000WO-JP01646.  
 XX 19-MAR-1999; 99JP-0075305.  
 XX 28-OCT-1999; 99JP-0306623.  
 XX (KURO/) KUROKAWA K.  
 XX (FUSO/) FUSO PHARM IND LTD.  
 XX (MIYA/) MIYATA T.  
 XX Miyata T;  
 PI  
 DR WPI; 2000-611642/58.  
 DR P-PSDB; AAB24142.  
 XX  
 PT Evaluating renal function comprises assaying megin protein in  
 PT biological sample -  
 XX  
 PS Example 5; Page 62-66; 93pp; Japanese.  
 CC The present invention describes a method for evaluating renal function.  
 CC The method comprises assaying megin protein in biological sample. Also  
 CC described are: (1) use of an anti-megsin protein antibody for diagnosing  
 CC renal function; and (2) a kit for detecting megin protein comprising:  
 CC (a) anti-megsin protein antibody attached to solid magnetic particles;  
 CC (b) direct or indirect fixing for the antibody to the particles; and  
 CC (c) a magnet. The process is useful for evaluating renal function and  
 CC diagnosing renal disorders by assaying megin protein in biological  
 CC samples (preferably urine or blood). The process is reproducible and  
 CC gives accurate results. The present sequence encodes the human megin

CC protein, which is given in the exemplification of the present invention.

XX SQ Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;

Query Match 100.0%; Score 1143; DB 21; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTGCTGAGCAAAATGAGAGTTTTCCTCAACCTGTTCCAGAGAGATGAT 60  
DB 1 ATGGCCCTCCCTGCTGAGCAAAATGAGAGTTTTCCTCAACCTGTTCCAGAGAGATGAT 60  
QY 61 GACAATCAAGGAATGAAATGTTTCTTCTCTGAGCCCTTCTGAGCCCTGCGCC 120  
DB 61 GACAATCAAGGAATGAAATGTTTCTTCTCTGAGCCCTTCTGAGCCCTGCGCC 120  
QY 121 CTGGCCGCTGGCGCTCAGNATGACTCCCTCTCTCAGATTGATGATGATGAT 180  
DB 121 CTGGCCGCTGGCGCTCAGNATGACTCCCTCTCTCAGATTGATGATGATGATGAT 180  
QY 181 AACACTGGCTCAGGATATGAAACTCTTCTAATAGTCAGTCAAGGCTCCAGTCTCAATG 240  
DB 181 AACACTGGCTCAGGATATGAAACTCTTCTAATAGTCAGTCAAGGCTCCAGTCTCAATG 240  
QY 241 AAAAGAGTTTTTCTGATATAAATGATCCCAAGATATGATCTCAGCATTTGGAAT 300  
DB 241 AAAAGAGTTTTTCTGATATAAATGATCCCAAGATATGATCTCAGCATTTGGAAT 300  
QY 301 GGGCTTTTCTGAAAGGTATGGCTTCTAATAGTACATGATGATGATGATGATGATGAT 360  
DB 301 GGGCTTTTCTGAAAGGTATGGCTTCTAATAGTACATGATGATGATGATGATGATGAT 360  
QY 361 TTATAGATGCAAGTGGAGGAGTTCATCTTACCAATCATTTAGAGACATGAGGT 420  
DB 361 TTATAGATGCAAGTGGAGGAGTTCATCTTACCAATCATTTAGAGACATGAGGT 420  
QY 421 AATATTAATAGTGGTGAATAAATGAACATGGCAAAATCAAGAACGTGATGATGATGAT 480  
DB 421 AATATTAATAGTGGTGAATAAATGAACATGGCAAAATCAAGAACGTGATGATGATGAT 480  
QY 481 GGTGGCATAGCTCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
DB 481 GGTGGCATAGCTCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 541 TGGCAATCAGCTTCCAGAGAGCAAAATCAAAATGCAATTTCAAAATCTCCCAAGTGC 600  
DB 541 TGGCAATCAGCTTCCAGAGAGCAAAATCAAAATGCAATTTCAAAATCTCCCAAGTGC 600  
QY 601 TCTGGAGAGGAGTGGCGATGATGATCAGAACGGAATTCATTTCTGTTATGAG 660  
DB 601 TCTGGAGAGGAGTGGCGATGATGATCAGAACGGAATTCATTTCTGTTATGAG 660  
QY 661 GACCATCAATGAGATTTCTGAGCTCAGTCAATGATGATGATGATGATGATGATGATGATGAT 720  
DB 661 GACCATCAATGAGATTTCTGAGCTCAGTCAATGATGATGATGATGATGATGATGATGATGAT 720  
QY 721 CTGCTGAGATGATCTCTGAAATGAAACAAACTGACCTTTCAGATCTTAATGAA 780  
DB 721 CTGCTGAGATGATCTCTGAAATGAAACAAACTGACCTTTCAGATCTTAATGAA 780  
QY 781 TGGACCAATCAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 781 TGGACCAATCAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 841 ATAGAGAGATTTATGAATGAAACAATTTGAGAGCCCTAGGCTGGAAGATATCTTT 900  
DB 841 ATAGAGAGATTTATGAATGAAACAATTTGAGAGCCCTAGGCTGGAAGATATCTTT 900  
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTTCTTGGGGGGTCTGATGATATATCAAGG 960  
DB 901 GATGAATCCAAAGCAGATCTCTCTGGGATTTCTTGGGGGGTCTGATGATATATCAAGG 960  
QY 961 ATGATGCAAAATTTACATAGAGTCACTGAGAGGCAACGAGGCTACTGCTCCACA 1020

DB 961 ATGATGCAAAATTTACATAGAGTCACTGAGAGGCAACGAGGCTACTGCTGCCACA 1020  
QY 1021 GGAAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080  
DB 1021 GGAAGTAATATTGTAGAAAAGCAACTCCCTCAGTCCACGCTGTTTAGAGCTGACACCCA 1080  
QY 1081 TTCTATTTTCTATCAGAGAGATGACATCATCTTATCAGTGGCAAGTTTCTTGCCT 1140  
DB 1081 TTCTATTTTCTATCAGAGAGATGACATCATCTTATCAGTGGCAAGTTTCTTGCCT 1140  
QY 1141 TGA 1143  
DB 1141 TGA 1143  
RESULT 3  
AAP82438 ID AAP82438 standard; cDNA; 1143 BP.  
XX AAP82438;  
AC AC  
XX  
DT 10-JUL-2001 (first entry)  
XX Human megilin nucleotide sequence.  
DE Human; megilin; mesangial cell proliferative nephritis; nephrotropic;  
KW transgenic mouse; glomerular disease; animal model; drug screening; ss.  
XX Homo sapiens.  
OS  
XX  
PH Key Location/Qualifiers  
FT CDS 1..1143  
FT /tag= a  
FT /product= "megilin"  
XX WO200124628-A1.  
XX 12-APR-2001.  
XX  
XX 06-OCT-2000; 2000WO-JP06988.  
XX  
XX 06-OCT-1999; 99JP-0285736.  
XX  
XX (KURO/) KUROKAWA K.  
XX (MIYA/) MIYATA I.  
XX Miyata T;  
XX  
XX WPI; 2001-300136/31.  
XX P-PSDB; AAB83075.  
XX  
XX Mouse model for mesangial cell proliferative nephritis for development and screening of new treatments -  
XX  
XX Claim 2; Page 42-44; 62pp; Japanese.  
XX  
XX The present sequence encodes human megilin. The human megilin coding sequence may be introduced into a mouse to produce an animal model of mesangial cell proliferative nephritis. The symptoms include enlargement of the mesangial base region, sedimentation of an immune complex and an increase in mesangial cells. The animal model is useful for analysing the pathology of chronic glomerular diseases and for screening compositions for prevention and treatment of the diseases.  
XX Highly uniform models can be made easily and in large numbers using this method.  
XX  
XX Sequence 1143 BP; 336 A; 230 C; 261 G; 316 T; 0 other;  
Query Match 100.0%; Score 1143; DB 22; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCTCCCTTCGTCGAGGAAATGAGAGTTTTCCTTCAACCTGTTTCAGAGATGGAT	60	DB	1081	TTCCCTATTGTTATCAGGAGGATGACATCATCTATTTCAGTGCAGAAAGTTTCTTGCCT	1140
DB	1	ATGGCTCCCTTCGTCGAGGAAATGAGAGTTTTCCTTCAACCTGTTTCAGAGATGGAT	60	QY	1141	TGA	1143
QY	61	GACAAATCAAGGAAATGGAATGTTCTTTTCTCTCTCTGAGCCTCTTGCCTGGCC	120	DB	1141	TGA	1143
DB	61	GACAAATCAAGGAAATGGAATGTTCTTTTCTCTCTCTGAGCCTCTTGCCTGGCC	120	RESULT 4			
QY	121	CTGGTCCGCTTGGGCTCAAGATGACTCCCTCTCTCAGATGATTAAGTTGCTTCATGTT	180	AAQ56670			
DB	121	CTGGTCCGCTTGGGCTCAAGATGACTCCCTCTCTCAGATGATTAAGTTGCTTCATGTT	180	ID	AAQ56670	standard; cDNA; 1950 BP.	
QY	181	AACACTGCTCAGGATATGGAATCTTCTTAATAGTCTAGTCAGGGCTCAGTCTCAACTG	240	XX	AC	AAQ56670;	
DB	181	AACACTGCTCAGGATATGGAATCTTCTTAATAGTCTAGTCAGGGCTCAGTCTCAACTG	240	XX	XX	25-MAR-2003 (updated)	
QY	241	AAAGAGTTTTTCTGATATAATGATCCCAAGGATATGATCTCAGATTTGTAAT	300	DT	16-AUG-1994	(first entry)	
DB	241	AAAGAGTTTTTCTGATATAATGATCCCAAGGATATGATCTCAGATTTGTAAT	300	XX	Human megakaryocyte differentiation factor coding sequence.		
QY	301	GGGCTTTTCTGAAAAAGTGATGCTTTTCAATAGGACTACATGAGTGGCCGAAAAA	360	XX	Human megakaryocyte differentiation factor; MDF; thrombopoietin;		
DB	301	GGGCTTTTCTGAAAAAGTGATGCTTTTCAATAGGACTACATGAGTGGCCGAAAAA	360	KW	haematopoietic stimulating factor; thrombocytopoiesis; platelet;		
QY	361	TTATACGATCCCAAGTGGAGGTTGACTTTTACGATATTTAGACACACTAGACGT	420	KW	bone marrow transplantation; cancer chemotherapy; ds.		
DB	361	TTATACGATCCCAAGTGGAGGTTGACTTTTACGATATTTAGACACACTAGACGT	420	XX	Homo sapiens.		
QY	421	AATATTAAATGAGGTTGAAATGAAACACATGGAATCAAGAACTGATTTGGTGAA	480	OS			
DB	421	AATATTAAATGAGGTTGAAATGAAACACATGGAATCAAGAACTGATTTGGTGAA	480	XX	Key	Location/Qualifiers	
QY	481	GTTGCTATAGCTATCTGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	540	PH	74..1216		
DB	481	GTTGCTATAGCTATCTGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	540	FT	/*tag= a		
QY	541	TGGCAATCAGCTTCCCAAGAGGAAACCAATTAATTCGCAATTTCAAACTCCCAAGTGC	600	FT	/product= megakaryocyte_differentiation_factor		
DB	541	TGGCAATCAGCTTCCCAAGAGGAAACCAATTAATTCGCAATTTCAAACTCCCAAGTGC	600	XX	EP583884-A1.		
QY	601	TCTGGAAGGAGCTGCCATGATGATCAGGAACGGAAGTTCAATTTGCTGTTATGAG	660	PN	23-FEB-1994.		
DB	601	TCTGGAAGGAGCTGCCATGATGATCAGGAACGGAAGTTCAATTTGCTGTTATGAG	660	PD	19-JUL-1993; 93EP-0305654.		
QY	661	GACCATCAATGAGGATTTCTGAGTCAATGATGATGATGATGATGATGATGATGATG	720	PF	17-JUL-1992; 92JP-0212305.		
DB	661	GACCATCAATGAGGATTTCTGAGTCAATGATGATGATGATGATGATGATGATGATG	720	PR	04-MAR-1993; 93JP-0067339.		
QY	721	CTGGCTGAGATGACCTCTCTGAAATGAAACAACTGACCTTTCAGATCTAATGAA	780	XX	(SUNR ) SUNTORY LTD.		
DB	721	CTGGCTGAGATGACCTCTCTGAAATGAAACAACTGACCTTTCAGATCTAATGAA	780	PA	(TSUJ/) TSUJIMOTO M.		
QY	781	TGGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTATTTTCTCAGTCAAG	840	PI	Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;		
DB	781	TGGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTATTTTCTCAGTCAAG	840	PI	Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamaichi K;		
QY	841	ATAGAGAGATTTATGAATGAAATATTTGAGAGCCCTAGGCTGGAAGATCTTT	900	XX	WPI; 1994-058782/08.		
DB	841	ATAGAGAGATTTATGAATGAAATATTTGAGAGCCCTAGGCTGGAAGATCTTT	900	DR	P-FSDB; AAR48379.		
QY	901	GATGATCCAAAGCAGATCTCTCTGGGATGCTTGGGGGCTGCTGTTATATATCAAGG	960	XX	New megakaryocyte differentiation factor - isolated from human		
DB	901	GATGATCCAAAGCAGATCTCTCTGGGATGCTTGGGGGCTGCTGTTATATATCAAGG	960	PT	epidermoid carcinoma cells, used to treat conditions involving a		
QY	961	ATGATGCAAAATTTTACATAGAGTCACTGAGAGGCGACCGAGGCTACTGCTGCCACA	1020	XX	decrease in platelets		
DB	961	ATGATGCAAAATTTTACATAGAGTCACTGAGAGGCGACCGAGGCTACTGCTGCCACA	1020	PS	Claim 13; Page 30-32; 47pp; English.		
QY	1021	GGAGTAAATTTTATGAAGAGCACTCCCTCAGTCCAGCTGTTTATGAGTACACCCA	1080	CC	The cDNA coding for human MDF was isolated from a library prepared		
DB	1021	GGAGTAAATTTTATGAAGAGCACTCCCTCAGTCCAGCTGTTTATGAGTACACCCA	1080	CC	using mRNA derived from human epidermoid carcinoma A431 cells. Human		
QY	1081	TTCCCTATTGTTATCAGGAGGATGACATCATCTATTTCAGTGCAGAAAGTTTCTTGCCT	1140	CC	MDF stimulates differentiation of megakaryocytes from myeloid cells		
DB	1081	TTCCCTATTGTTATCAGGAGGATGACATCATCTATTTCAGTGCAGAAAGTTTCTTGCCT	1140	CC	in the presence of IL-3. The MDF acts in vivo as a thrombopoietin		
QY				CC	making it useful for treatment of diseases involving a decrease in		
DB				CC	platelet number (esp. thrombocytopoiesis) such as occurs in bone		
QY				CC	marrow transplantation and in chemotherapy.		
DB				CC	(Updated on 25-MAR-2003 to correct PN field.)		
QY				XX	Sequence 1950 BP; 570 A; 384 C; 407 G; 589 T; 0 other;		
DB				XX	Query Match 100.0%; Score 1143; DB 15; Length 1950;		
QY				XX	Best Local Similarity 100.0%; Pred. No. 0;		
DB				XX	Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY				XX	1 ATGGCTCCCTTCGTCGAGGAAATGAGAGTTTTCCTTCAACCTGTTTCAGAGATGGAT	60	
DB				XX	74 ATGGCTCCCTTCGTCGAGGAAATGAGAGTTTTCCTTCAACCTGTTTCAGAGATGGAT	133	



QY	61	GACAATCAGGAAATGGAAATGTTCTTTTCCTCTCTGAGCCTCTTTTCGCTGCCCTCGGCC	120
Db	134	GACAATCAGGAAATGGAAATGTTCTTTTCCTCTCTGAGCCTCTTTTCGCTGCCCTCGGCC	193
QY	121	CTGCTCGGCTTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGCTTCATGTT	180
Db	134	CTGCTCGGCTTGGGCGCTCAGATGACTCCCTCTCTCAGATTGATAGTTGCTTCATGTT	253
QY	161	AACACTGCCCTCAGGATATGGAACTCTTCTAATAAGTCACTCAGGCTCCAGCTCTCAACTG	240
Db	254	AACACTGCCCTCAGGATATGGAACTCTTCTAATAAGTCACTCAGGCTCCAGCTCTCAACTG	313
QY	241	AAAAAGAGTTTTTTCGATATATAAATGCATGCCCAAGAATTAATGATCTCAGCATTTGTAAT	300
Db	314	AAAAAGAGTTTTTTCGATATATAAATGCATGCCCAAGAATTAATGATCTCAGCATTTGTAAT	373
QY	301	GGGCTTTTGTGCTGAAAAAGTGTATGCTTTTCATAAGGACTACATTGAGTGTGCCGAAAAA	360
Db	374	GGGCTTTTGTGCTGAAAAAGTGTATGCTTTTCATAAGGACTACATTGAGTGTGCCGAAAAA	433
QY	361	TTATACGATGCCAAGTTCGAGCGGTTGACTTTTACGATTCATTTAGAACACATPAGACGT	420
Db	434	TTATACGATGCCAAGTTCGAGCGGTTGACTTTTACGATTCATTTAGAACACATPAGACGT	493
QY	421	AATTAATAATTAAGTGGGTGMAAATGAACAACATGGCAAATCAAGAAACGTGATGGTGA	480
Db	494	AATTAATAATTAAGTGGGTGMAAATGAACAACATGGCAAATCAAGAAACGTGATGGTGA	553
QY	481	GGTGGCATTAAGCTCACTGCTGTATATGGTGCTGGTGAATGCTGTGTAATCAAAAGGCAAG	540
Db	554	GGTGGCATTAAGCTCACTGCTGTATATGGTGCTGGTGAATGCTGTGTAATCAAAAGGCAAG	613
QY	541	TGGCAATCAGGCTTTCACCAAGAGCGAAACCATAAATTCGCAATTTCAAACTCCCAAGTGC	600
Db	614	TGGCAATCAGGCTTTCACCAAGAGCGAAACCATAAATTTCCCAATTTCCCAAGTGC	673
QY	601	TTTGGGAGGCGATGCGCAATGATCATCAGGAACCGGAAGTTCAATTTGTCTGTTATTGAG	660
Db	674	TTTGGGAGGCGATGCGCAATGATCATCAGGAACCGGAAGTTCAATTTGTCTGTTATTGAG	733
QY	661	GAOCCATCAATGAGATCTCTCAGCTCGATACATATGTTGGCATTAACATGATGATGTTCTG	720
Db	734	GACCCATCAATGAGATCTCTCAGCTCGATACATATGTTGGCATTAACATGATGATGTTCTG	793
QY	721	CTCCCTGAGATGACCTCTCGAAATTTGAAACCAAACTTGACCTTTCAAGAACTTAATGGAA	780
Db	794	CTCCCTGAGATGACCTCTCGAAATTTGAAACCAAACTTGACCTTTCAAGAACTTAATGGAA	853
QY	781	TGGACCAATCCAGGCGAATGACCTCTTAAGTATGTTGAGTATTTTTCCTCAGTTCAG	840
Db	854	TGGACCAATCCAGGCGAATGACCTCTTAAGTATGTTGAGTATTTTTCCTCAGTTCAG	913
QY	841	ATAGAGAAATATTAATGAAATGAACAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTT	900
Db	914	ATAGAGAAATATTAATGAAATGAACAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTT	973
QY	901	GATGAATCCAAAGCAGATCTCTCTGGGANTGCTTCGGGGGGTCTGTATATATCAAGG	960
Db	974	GATGAATCCAAAGCAGATCTCTCTGGGANTGCTTCGGGGGGTCTGTATATATCAAGG	1033
QY	961	ATGATGCAAAATCTTACATAGAGTCACTCAGGAGGCGACCGGCTACTGCTGCCACA	1020
Db	1034	ATGATGCAAAATCTTACATAGAGTCACTCAGGAGGCGACCGGCTACTGCTGCCACA	1093
QY	1021	GGAGTATATTTGAGAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTTGACCCACA	1090
Db	1094	GGAGTATATTTGAGAAAGCAACTCCCTCAGTCCAGCTGTTTAGAGCTTGACCCACA	1153
QY	1081	TTCCATTTGTTATCAGGAGGATGACATCATTTATTCAGTGCAGAAAGTTTCTGCCCT	1140
Db	1154	TTCCATTTGTTATCAGGAGGATGACATCATTTATTCAGTGCAGAAAGTTTCTGCCCT	1213
QY	1141	TGA	1143

Db 1214 TGA 1216

||||

RESULT 5

ABX76260

ID ABX76260 standard; DNA; 2249 BP.

AC ABX76260;

XX

DT 02-APR-2003 (first entry)

XX

DE lung cancer-associated polynucleotide #126.

XX

DE lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;

KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX

OS Unidentified.

XX

FN WO200286443-A2.

XX

PD 31-OCT-2002.

XX

PF 18-APR-2002; 2002WO-US12476.

XX

PR 18-APR-2001; 2001US-284770P.

XX

PR 10-MAY-2001; 2001US-290492P.

PR 09-NOV-2001; 2001US-339245P.

PR 13-NOV-2001; 2001US-350656P.

PR 29-NOV-2001; 2001US-334370P.

PR 12-APR-2002; 2002US-372246P.

XX

XX (EOSB-) EOS BIOTECHNOLOGY INC.

PA

PA Aziz N, Murray R;

XX

PI WPI; 2003-093161/08.

DR

DR P-PSDB; ABUS6533.

XX

PT Detecting a lung cancer-associated transcript in a cell from a patient

PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased

PT expression in lung cancer -

XX

PS Claim 22; Page 288; 453pp; English.

XX

CC The invention relates to a method for detecting a lung cancer-associated

CC transcript in a cell from a patient, comprising contacting a biological

CC sample from the patient with a polynucleotide that selectively hybridises

CC to a sequence that is at least 80 % identical to a gene that exhibits

CC increased or decreased expression in lung cancer samples. Lung

CC cancer-associated polynucleotides and polypeptides are used for

CC identifying a compound that modulates a lung cancer-associated

CC polypeptide, for inhibiting proliferation of a lung cancer-associated

CC cell to treat lung cancer in a patient and for treating a mammal having

CC lung cancer by administering a modulatory compound identified. The

CC methods are useful for treating lung cancer, such as small cell lung

CC cancer, non-small cell lung cancer or other benign or precancerous

CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive

CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial

CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides

CC and polypeptides are useful for diagnostic purposes and as targets for

CC screening for therapeutic compounds that modulate lung cancer, such as

CC antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated

CC polynucleotides of the invention.

XX

SQ Sequence 2249 BP; 649 A; 452 C; 461 G; 693 T; 4 other;

XX

Query Match 100.0%; Score 1143; DB 25; Length 2249;

Best Local Similarity 100.0%; Pred No. 0;

Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCCCTCCCTCTCGACAAATGCGAGCTTTTCCTTCAACCTGTTTCAGAGAGATGGAT	60
Db	365	ATGGCCCTCCCTCTCGACAAATGCGAGCTTTTCCTTCAACCTGTTTCAGAGAGATGGAT	424
QY	61	GACAAATCAAGGAATGGAATGTTCTTTTCTCTCTGAGCCCTTCGCTGCCCTGGCC	120
Db	425	GACAAATCAAGGAATGGAATGTTCTTTTCTCTCTGAGCCCTTCGCTGCCCTGGCC	484
QY	121	CTGGTCCTCTGGGCGCTCAAGATGACTCCCTCTCTCAAGATGATGATGTTCTCATGTT	180
Db	485	CTGGTCCTCTGGGCGCTCAAGATGACTCCCTCTCTCAAGATGATGATGTTCTCATGTT	544
QY	181	RACATGCTCAGGATATGGAATCTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG	240
Db	545	AACATCTGCTCAGGATATGGAATCTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG	604
QY	241	AAAGAGTTTTTCTGATATAATGCAATCCCAAGGATTAATGATCTCAGCATTTGAAAT	300
Db	605	AAAGAGTTTTTCTGATATAATGCAATCCCAAGGATTAATGATCTCAGCATTTGAAAT	664
QY	301	GGGCTTTTGTGAAAAGTGTATGGCTTTCAATAGGACTACATGAGTGTCCGAAAAA	360
Db	665	GGGCTTTTGTGAAAAGTGTATGGCTTTTCAATAGGACTACATGAGTGTCCGAAAAA	724
QY	361	TTATACGATGCGAAAGTGGAGCGAGTTGACTTTTACGAATCATTTAGAGCACTAGAGCT	420
Db	725	TTATACGATGCGAAAGTGGAGCGAGTTGACTTTTACGAATCATTTAGAGCACTAGAGCT	784
QY	421	AAATTAATTAAGTGGSTTGAAATGAACACATGCGCAAAATCAAGAACGTGATGTTGAA	480
Db	785	AAATTAATTAAGTGGSTTGAAATGAACACATGCGCAAAATCAAGAACGTGATGTTGAA	844
QY	481	GGTGCAATAGCTCATCTCTCTTAATGCTGCTGTAATGCTGTCTCTCAAGGCAAG	540
Db	845	GGTGCAATAGCTCATCTCTCTTAATGCTGCTGTAATGCTGTCTCTCAAGGCAAG	904
QY	541	TGGCAATCAGCTTTACCAAGAGCGAAACCAATTTGCCATTTCAAACTCCCAAGTGC	600
Db	905	TGGCAATCAGCTTTACCAAGAGCGAAACCAATTTGCCATTTCAAACTCCCAAGTGC	964
QY	601	TTGGGAAGGCGTCCCATGATGCAATCAGGAACGGAAGTTCAATTTGCTCTTATTGAG	660
Db	965	TTGGGAAGGCGTCCCATGATGCAATCAGGAACGGAAGTTCAATTTGCTCTTATTGAG	1024
QY	661	GACCAATCAATGAAGATCTTGGAGTGATGATCAATGTTGGCATTAACATGACCTTCTG	720
Db	1025	GACCAATCAATGAAGATCTTGGAGTGATGATCAATGTTGGCATTAACATGACCTTCTG	1084
QY	721	CTGGCTGAGATGACCTCTCTGAAATGAAACCAAACTGACCTTTGAGATCTAATGAA	780
Db	1085	CTGGCTGAGATGACCTCTCTGAAATGAAACCAAACTGACCTTTGAGATCTAATGAA	1144
QY	781	TGGACCAATCAAGGCGAATGACCTCTAATGATGTTGAGTATTTTTCCTCAGTTCAG	840
Db	1145	TGGACCAATCAAGGCGAATGACCTCTAATGATGTTGAGTATTTTTCCTCAGTTCAG	1204
QY	841	ATAGAGAGATTTATGAAATGAAACATATTTGAGAGCCCTAGGCTGGAAGATATCTTT	900
Db	1205	ATAGAGAGATTTATGAAATGAAACATATTTGAGAGCCCTAGGCTGGAAGATATCTTT	1264
QY	901	GATGAATCCAAAGCATCTCTCTGGGATTTGCTTGGGGGCTCTCTGTATATATCAAGG	960
Db	1265	GATGAATCCAAAGCATCTCTCTGGGATTTGCTTGGGGGCTCTCTGTATATATCAAGG	1324
QY	961	ATGATGCAAAATTTTACATAGAGTCACTGAGGAGGCGACCGAGGCTACTCTGCCACA	1020
Db	1325	ATGATGCAAAATTTTACATAGAGTCACTGAGGAGGCGACCGAGGCTACTCTGCCACA	1384
QY	1021	GGAGTAATATTGAGAAAGCACTCCCTCAGTCCAGCTGTTTGTAGAGTCAACCCCA	1080
Db	1385	GGAGTAATATTGAGAAAGCACTCCCTCAGTCCAGCTGTTTGTAGAGTCAACCCCA	1444

QY 1081 TTCTATTTCATCAGGAGGATGACATCATCTTATTCAGTGGCAAACTTCTTCCCT 1140

Db 1445 TTCTATTTCATCAGGAGGATGACATCATCTTATTCAGTGGCAAACTTCTTCCCT 1504

QY 1141 TGA 1143

Db 1505 TGA 1507

RESULT 6

AAH48181

ID AAH48181 standard; DNA; 1229 BP.

XX AC AAH48181;

XX XX

DT 21-SEP-2001 (first entry)

XX Rat megin protein coding sequence.

DE DE

XX Rat; megin; renal mesangial cell; mesangium proliferative nephritis; ds.

XX OS Rattus norvegicus.

XX PN WO200148019-A1.

XX PD 05-JUL-2001.

XX PF 26-DEC-2000; 2000WO-JF09251.

XX PR 28-DEC-1999; 99JP-0373677.

XX PA (KURO/) KUROKAWA K.

PA (MIYA/) MIYATA I.

XX MIyata T;

XX WPI; 2001-425651/45.

DR P-PSDB; AAG64286.

XX

XX New antibody recognizing a partial sequence of rat megin protein for diagnosis of mesangium proliferative nephritis -

XX Example 5; Page 52-54; 63pp; Japanese.

XX The present invention relates to a novel antibody which recognises a peptide consisting of residues 341-354 of rat megin protein. The present sequence is the coding sequence for rat megin, which was used in the present invention. Megin is highly expressed in renal mesangial cells and its level is elevated in mesangium proliferative nephritis. Assay of the serum or urine level using the antibody is therefore indicative of this type of disorder.

XX Sequence 1229 BP; 378 A; 243 C; 276 G; 332 T; 0 other;

QY 1 ATGGGCTCCCTGCTGTCAGCAAAATGAGATTTTGTCTCAACCTGTTTCAGAGATGGAT 60

Db 8 ATGGGCTCCCTGCTGTCAGCAAAATGAGATTTTGTCTCAACCTGTTTCAGAGATGGAT 67

QY 61 GACATCAAGGAATGGAATGTTCTTTCTCTCTCTGAGCCCTTCGCTGCCCTGGCC 120

Db 68 AGTAGTCAAGGAACGGAATGTTATTTCTCTCTCTGAGCCCTTCGCTGCCCTGGCC 127

QY 121 CTGGTCCCTGCGCTCAAGTCACTCCCTCTCTCAGATTGATGATGTTCTTCTT 180

Db 128 CTAAATCCCTTTGGTGGTCTCGAGTCACTGTGACGTCAAGATTGCAAGCCCTGCACTTT 187

QY 181 AACACTGCTCAGGATATGGAATCTTCTTAATAGTCAAGGCTCCAGTCTCAACTG 240



```
QY 421 AATAATTAATAGCTGGTTGAAATGAACACATGCGAAATCAAGAACGTGATGGTGAA 480
Db 428 AAAATTAATTAATGGAATGAATGAACACATGCGAAATCAAGAACGTGATGGTGAA 487
QY 481 GGTGGCATTAAGCTCATCTGCTGTATTAATGGTGGTGAATGCTGTGATCTTCAAAGGCAAG 540
Db 488 ACCAGCTCAGCTCATCAGCTGTGATGGTGGTGAATGCTGTGATCTTCAAAGGCAAG 547
QY 541 TGGCAATCAGCTTCCAGAGCGAAGCAATTAATGCGAATTTCAAATCTCCCAAGTGC 600
Db 548 TGGAAATCGGCTTCCAGAGCGAAGCAATTAATGCGAATTTCAAATCTCCCAAGTGC 607
QY 601 TCTGGGAAGCGCTGCCATGATGATCAGAGCAAGCAATTAATGCGAATTTCAAATCTCCCAAGTGC 660
Db 608 CTTGGAAAGCAATTAATGATGATCAGAGCAAGCAATTAATGCGAATTTCAAATCTCCCAAGTGC 667
QY 661 GACCATCAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 720
Db 668 GAGCCCAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 727
QY 721 CTGCTGAGATGATCTTGAATTAATGCGAATTTCAAATCTCCCAAGTGC 780
Db 728 TTGCGCGAGATGATCTTGAATTAATGCGAATTTCAAATCTCCCAAGTGC 787
QY 781 TGGCAATCAGCTTCCAGAGCGAAGCAATTAATGCGAATTTCAAATCTCCCAAGTGC 840
Db 788 TGGCAATCAGCTTCCAGAGCGAAGCAATTAATGCGAATTTCAAATCTCCCAAGTGC 847
QY 841 ATAGAGAAATTAATGAATGAAGCAATTAATGCGAATTTCAAATCTCCCAAGTGC 900
Db 848 ATAGAGAAATTAATGAATGAAGCAATTAATGCGAATTTCAAATCTCCCAAGTGC 907
QY 901 GATGATCAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 960
Db 908 GTTGAATCAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 967
QY 961 ATGATGCAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 1020
Db 968 CTAATGCAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 1027
QY 1021 GGAATTAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 1080
Db 1028 GGAATTAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 1087
QY 1081 TTCCATTAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 1140
Db 1088 TTCCATTAATTAAGATCTTGGATCAGATCAATTAATGCGAATTTCAAATCTCCCAAGTGC 1147
QY 1141 TGA 1143
Db 1148 TGA 1150
```

## RESULT 8

AAAC5238  
ID AAC5238 standard; DNA; 1229 BP.

AC AAC5238;

XX 30-JAN-2001 (first entry)

DE Rat meginin nucleotide sequence SEQ ID NO:18.

KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;  
KW IgA; immunoglobulin A; detection; renal function; renal disorder;  
KW diagnosis; biological sample; blood; urine; ss.

OS Rattus norvegicus.

XX W0200057189-A1.

PN 28-SEP-2000.

PD

XX 17-MAR-2000; 2000WO-JP01646.  
XX 19-MAR-1999; 99JP-0075305.  
PR 28-OCT-1999; 99JP-0306623.  
XX (KURO/) KUROKAWA K.  
PA (FUSO) FUSO PHARM. IND. LTD.  
PA (MIYA) MIYATA T.  
XX Miyata T;  
XX WPI: 2000-611642/58.  
DR P-PSDB; AAB24150.  
XX Evaluating renal function comprises assaying meginin protein in  
PT biological sample -  
XX Disclosure; Page 77-81, 93pp; Japanese.  
XX The present invention describes a method for evaluating renal function.  
CC The method comprises assaying meginin protein in biological sample. Also  
CC described are: (1) use of an antimeginin protein antibody for diagnosing  
CC renal function; and (2) a kit for detecting meginin protein comprising:  
CC (a) antimeginin protein antibody attached to solid magnetic particles;  
CC (b) direct or indirect fixing for the antibody to the particles; and  
CC (c) a magnet. The process is useful for evaluating renal function and  
CC diagnosing renal disorders by assaying meginin protein in biological  
CC samples (preferably urine or blood). The process is reproducible and  
CC gives accurate results. The present sequence encodes the rat meginin  
CC protein, which is given in the exemplification of the present invention.  
XX Sequence 1229 BP; 377 A; 242 C; 274 G; 330 T; 6 other;

Query Match 65.9%; Score 753; DB 21; Length 1229;

Best Local Similarity 78.5%; Pred. No. 1.5e-216;

Matches 897; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

```
QY 1 ATGGGCTCCCTTGTCTGCAACAAATGCAAGATTTTGGTTCAACCTGTTTCAGAGAGATGAT 60
Db 8 ATGGGCTCCCTTGTCTGCAACAAATGCAAGATTTTGGTTCAACCTGTTTCAGAGAGATGAT 67
QY 61 GACATCAAGAAATGAAATGTTTCTTCTCTGAGCTCTTCTGCTCGCTCGCTCGCTCGCT 120
Db 68 ACTAGTCAAGAAACGGAATGATTTCTTCTCTGAGCTCTTCTGCTCGCTCGCTCGCTCGCT 127
QY 121 CTGGTCCCTTGGGCGCTCAAGTGAATCTCTCTCTGAGTGAATGATGATGATGATGATGAT 180
Db 128 CTAATCCGTTTGGTGTCTGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 187
QY 181 AACCTGCTCAGGATATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 188 ATCTCCCATCAGACAAAGGAATTCATGAAACAGTCAAGTCAAGTCAAGTCAAGTCAAGT 247
QY 241 AAAAGAGTTTCTGATATAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 248 AAAAGAGTTTCTGATATAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 307
QY 301 GGCCTTTTGTGAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 308 GGAATTTTGTGAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 367
QY 361 TTATAGATGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 420
Db 368 TTATAGATGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 427
QY 421 AATATTAAATAGTGGTGTGAAAATGAAACACATGCGAAAATCAAGACGATGATGATGATGAT 480
Db 428 AATATTAAATAGTGGTGTGAAAATGAAACACATGCGAAAATCAAGACGATGATGATGATGAT 487
QY 481 GGTGGCATAGCTCATCTCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 488 AGCAGCTCAGCTCATCTCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 547
```



```

QY 661 GACCCATCAATGAAGATCTTTGAGCTCAGATACATAATGGTGGCATAAACAATGATGCTTCG 720
Db 668 GAGCCACCAATGAGATCTTTGAGCTACAAATATCATGGTGGCATAAAGCATGTATCATATG 727
QY 721 CTGCTCAGAGTACCTCTCTGAAATGAAACCAACTGACCTTTGAGAACTCAATCGAA 780
Db 728 TTGCCCCAGAGTGAOCTATCCGAAATGAAAGCAAGCTGAGTTCCAGAACTCAATGGAC 787
QY 781 TGGACCAATCCAAAGCGAATGACCTCTTAAGTATGTTGAGTATTTTCTCAGTTCAG 840
Db 788 TGGCAATAGCAGGAAGATGAATCTCAGTATGATGTGTTTCTCCCCAGTTCAG 847
QY 841 ATAGAGAAATATGAATGAAGAATATTTGAGAGCCCTAGGGCTGAAAGATATCTTT 900
Db 848 ATAGAGAAATATGAATGAGAGCACTTGAATCTGTAGGCTTGAAGACATCTTT 907
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATGCTTCGGGGGTGCTGTGTATATATCAAG 960
Db 908 GTTGAATCCAGGGCTGATCTCTGGAATGTCCTCTGAGGTGCTCTATGTATCAAG 967
QY 961 ATGATGACAAATCTTACATAGAGTCACTGAGGAGGCGACCGAGGCTACTCTGCCACA 1020
Db 968 CTAATGACCAAGTCCCTCATAGAGTCTCAGAGGAAGCGACCGAGGCAACTCTGCCACA 1027
QY 1021 GGAAGTAAATTTAGAAAGAACTCCCTCAGTCAAGCTGTCAGCTGTTAGAGCTGACACCA 1080
Db 1028 GAAAGTAACTATCTGAAAGCTATCTCTGAAATCCAGGTGTTGAGAGCTGACGCCCC 1087
QY 1081 TTCTTATTTGTTATCAGGAAGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT 1140
Db 1088 TTCTGTTTGTATAGGAAGATGGCATCACTTATTTACTGCAAGTCTCTGTGCT 1147
QY 1141 TGA 1143
Db 1148 TGA 1150

RESULT 10
AA56714
ID AA56714 standard; DNA; 1147 BP.
XX
AC AA56714;
XX
DT 14-JUL-1999 (first entry)
XX
DE Mouse megisin DNA.
XX
KW Megisin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
KW human; rat; murine; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
CDS 1..1107
ET /*tag= a
ET /product= "megisin"
FT /note= "Partial sequence, no start codon given"
XX
PN W09915652-A1.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WC-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
W21; 1999-276983/23.

```

```

DR P-PSDB; AA508256.
XX
PT Megsin protein expressed specifically in mesangial cells
XX
PS Claim 4; Page 72-76; 100pp; Japanese.
XX
CC This invention describes the isolation of novel megisin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as IgA nephropathy.
XX
SQ Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;

Query Match 65.5%; Score 748.6; DB 20; Length 1147;
Best Local Similarity 79.8%; Pred. No. 3.2e-215;
Matches 883; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

QY 37 TTCACCTGTTTCAGAGATGATGATCAATCAGAAATGGAATGTTCTTTCTCT 96
Db 1 TTCGATTTATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 97 CTGAGCCTCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156
Db 61 CTGAGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 157 CAGATTGATAGTGTCTTCATGTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 216
Db 121 CAGATTGCAAGGCACTGCTGCTTAACTACCATCAAGCAAGGAATCATCTTAAT 180
QY 217 CAGTCAGGCTCCAGTCTCACTGCAAGGAAGATTTTCTGATATAAATGATCCCAAG 276
Db 181 CAGCAGAGCTCTGATGATCAATGCAAGGAATTTCTGCTGATCAATCACTCAAT 240
QY 277 GATTATGATCTCAGCATCTGAATGGGCTTTTGTGAAAAGTGTATGCTTTCTAAG 336
Db 241 GATTATGAACTCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 337 GACTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
Db 301 AACTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 397 AATCATTTTGAAGCACTAGACGTAATTAATTAATGCTGCTGCTGCTGCTGCT 456
Db 361 AATGATGCAAGATCAGATTTAAATTAATTAATGATGATGATGATGATGATGAT 420
QY 457 AAAATCAAGAACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
Db 421 AAGATCAAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 517 AATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
Db 481 AAGCTGTGTTTACTTCAAGGCAATGGAATGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 577 TSCCATTTTCAATCTCCCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
Db 541 TSCCGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 637 AAGTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db 601 CGGTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 697 GGTGGCATAPACATGTACGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
Db 661 GGTGGCATAPACATGTACATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 757 CTGACCTTTTCAATCTTAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 816
Db 721 CTGAGTTTCCAGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 817 GAGGTATTTTCTCAGTCTCAGATGAGAGAAATGGAATGGAATGGAATGGAAT 876
Db 781 AAGCTGTTTCTCCCCAGTTCAGATGAGAGAAATGGAATGGAATGGAATGGAAT 840

```

QY 877 GCCTAGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTTGGATGCTTCC 936  
 |||  
 Db 841 TCTTAGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTTGGATGCTTCC 900  
 |||  
 QY 937 GGGGTCGTCTGTATATCAAGATGATGCACAAATCTTACATAGAGTCACTGAGGAG 996  
 |||  
 Db 901 GGAGTCGTCTCTAGTATCAAGCTAATGACACAGTCAATCATAGAGTCTCAGAGGAG 960  
 |||  
 QY 997 GGCACCGAGCTACTGTCGCCACAGGAGTAATATTGTAGAAAGCAACTCCCTCAGTCC 1056  
 |||  
 Db 961 GGCACGAGGCTACTGTCGCCACAGGAGTAATATTGTAGAAAGCAACTCCCTCAGTCC 1020  
 |||  
 QY 1057 ACCCTGTTTAGAGCTGACCAACCATTCCTATTCTTATCAGGAAGGATGACATCACTTA 1116  
 |||  
 Db 1021 ACAGTGTTCAGAGCGACCGCCCTTCTGTTTGTCTATCAGGAGATGACATCACTTA 1080  
 |||  
 QY 1117 TTCACTGCAAGTCTTCTGCTTGA 1143  
 |||  
 Db 1081 TTCACTGCAAGTCTTCTGCTTGA 1107  
 |||

## RESULT 11

AAC55239

ID AAC55239 standard; DNA; 1147 BP.

AC AAC55239;

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

SQ Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;  
 Query Match 65.5%; Score 748.6; DB 21; Length 1147;  
 Best Local Similarity 79.8%; Pred. No. 3.2e-215;  
 Matches 883; Conservative 0; Mismatches 224; Indels 0; Gaps 0;  
 QY 37 TTCAAGCTCTTCAGAGAGTGGATGACATCAAGGAATGGAATGTCTTTCTCTCT 96  
 |||  
 Db 1 TTCAAGCTCTTCAGAGAGTGGATGACATCAAGGAATGGAATGTCTTTCTCTCT 60  
 |||  
 QY 97 CTGAGCTCTTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 156  
 |||  
 Db 61 CTGAGCATCTTCAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 120  
 |||  
 QY 157 CAGATTGAATGTTCT 216  
 |||  
 Db 121 CAGATTGAATGTTCT 180  
 |||  
 QY 217 CAGTCAGGCTCCAGTCTCACTGAAAGAGTCTTTCATATATAATGAATCCACAAG 276  
 |||  
 Db 181 CAGTCAGGCTCCAGTCTCACTGAAAGAGTCTTTCATATATAATGAATCCACAAG 240  
 |||  
 QY 277 GATTATGATCTCAGCATTTGTAATGGCTTTTCTGTAAGAGTGTATGGCTTTCTA 336  
 |||  
 Db 241 GATTATGATCTCAGCATTTGTAATGGCTTTTCTGTAAGAGTGTATGGCTTTCTA 300  
 |||  
 QY 337 GACTCAATTTGAGTGTGCGGAAAATTTATACGATCCAAAGTGGAGGAGTTGACTT 396  
 |||  
 Db 301 AACTACATTTGAGTGTGCGGAAAATTTATACGATCCAAAGTGGAGGAGTTGACTT 360  
 |||  
 QY 397 AATCATTATAGAGACACTAGACGTATATTAATTAAGTGGGTTGAAATGAACACATGG 456  
 |||  
 Db 361 AATGATGTCAGAGTACCGATTTAAATTAATTAATTAATTAATTAATTAATTAAT 420  
 |||  
 QY 457 AATATCAAGACGTTGTTGTTGAGGAGTGGCATAAGCTCATCTCTGTAATGGTCTG 516  
 |||  
 Db 421 AAGATCAAGAGGTTGTTGCGGACACAGCTCAGCTGCTGCTGCTGCTGCTGCTG 480  
 |||  
 QY 517 AATGCTGTCTACTTCAAGGCAAGTGGCAATCAGCCTTCAACAGAGGCAAAACCAAT 576  
 |||  
 Db 481 AATGCTGTCTACTTCAAGGCAAGTGGCAATCAGCCTTCAACAGAGGCAAAACCAAT 540  
 |||  
 QY 577 TGCCATTTCAAATCTCCCAAGTCTCTGCGGAGGAGTGGCCTGCTGCTGCTGCTGCT 636  
 |||  
 Db 541 TGCCATTTCAAATCTCCCAAGTCTCTGCGGAGGAGTGGCCTGCTGCTGCTGCTGCT 600  
 |||  
 QY 637 AAGTTCAATTTCTCTGTTATTGAGGAGCCCATCAATGAAGATCTTGAGCTCAGATCA 696  
 |||  
 Db 601 CGGTTCAATTTCTCTGTTATTGAGGAGCCCATCAATGAAGATCTTGAGCTCAGATCA 560  
 |||  
 QY 697 GGTGGCATTAACATGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756  
 |||  
 Db 661 GGTGGCATTAACATGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 |||  
 QY 757 CTGACCTTTCAAGATCTAATGAATGGACCAATCCAGGCGAATGACCTCTGAAATGAAA 816  
 |||  
 Db 721 CTGACCTTTCAAGATCTAATGAATGGACCAATCCAGGCGAATGACCTCTGAAATGAAA 780  
 |||  
 QY 817 GAGGTATTTTCTCTGTTATTGAGGAGCCCATCAATGAAGATCTTGAGCTCAGATCA 876  
 |||  
 Db 781 AAGCTGTTTCTCTGTTATTGAGGAGCCCATCAATGAAGATCTTGAGCTCAGATCA 840  
 |||  
 QY 877 GCCTTAGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTTGGATGCTTCC 936  
 |||  
 Db 841 TCTTAGGCTGAAGATATCTTTGATGATCCAAAGCAGATCTCTTGGATGCTTCC 900  
 |||  
 QY 937 GGGGTCGTCTGTATATCAAGATGATGCACAAATCTTACATAGAGTCACTGAGGAG 996  
 |||  
 Db 901 GGAGTCGTCTCTAGTATCAAGCTAATGACACAGTCAATCATAGAGTCTCAGAGGAG 960  
 |||  
 QY 997 GGCACCGAGCTACTGTCGCCACAGGAGTAATATTGTAGAAAGCAACTCCCTCAGTCC 1056  
 |||  
 Db 961 GGCACCGAGCTACTGTCGCCACAGGAGTAATATTGTAGAAAGCAACTCCCTCAGTCC 1020  
 |||

Evaluating renal function comprises assaying megin protein in biological sample

Disclosure; Page 85-88; 93pp; Japanese.

The present invention describes a method for evaluating renal function. The method comprises assaying megin protein in biological sample. Also described are: (1) use of an antimegin protein antibody for diagnosing renal function; and (2) a kit for detecting megin protein comprising: (a) antimegin protein antibody attached to solid magnetic particles; (b) direct or indirect fixing for the antibody to the particles; and (c) a magnet. The process is useful for evaluating renal function and diagnosing renal disorders by assaying megin protein in biological samples (preferably urine or blood). The process is reproducible and gives accurate results. The present sequence encodes the mouse megin protein, which is given in the exemplification of the present invention.

QY 1057 AGCTGTTTACAGCTGACACCCATTCCTATTGTTATATCAGGAGATGACATCATCTTA 1116  
 DB 1021 ACAGTGTTCAGAGCCGACGCGCCCTTTCTGTTGTCAACAAGAGATGACATCATCTTA 1080  
 QY 1117 TTCACTGGCAAGCTTCTTGCCCTTGA 1143  
 DB 1081 TTTACTGGCAAGCTTCTTGCCCTTGA 1107

RESULT 12  
 AAF82440  
 ID AAF82440 standard; cDNA; 1147 BP.  
 XX  
 AC AAF82440;  
 DT 10-JUL-2001 (first entry)  
 XX  
 DE Mouse megisin nucleotide sequence.  
 XX  
 KW Mouse; megisin; mesangial cell proliferative nephritis; nephrotropic;  
 KW transgenic mouse; glomerular disease; animal model; drug screening; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..1107  
 FT /\*tag= a  
 FT /product= "megisin"  
 XX  
 PN W0200124628-AL.  
 XX  
 PD 12-APR-2001.  
 XX  
 PF 06-OCT-2000; 2000WO-JP06988.  
 XX  
 PR 06-OCT-1999; 99JP-0285736.  
 XX  
 PA (KURO/) KUROKAWA K.  
 PA (MIYA/) MIYATA T.  
 XX  
 PI Miyata T;  
 XX  
 DR WPI; 2001-300136/31.  
 DR P-PSDB; AAB83077.  
 XX  
 PT Mouse model for mesangial cell proliferative nephritis for development  
 PT and screening of new treatments  
 XX  
 PS Disclosure; Page 50-52; 62pp; Japanese.  
 XX  
 CC The present sequence encodes mouse megisin. The human megisin coding  
 CC sequence may be introduced into a mouse to produce an animal model of  
 CC mesangial cell proliferative nephritis. The symptoms include  
 CC enlargement of the mesangial base region, sedimentation of an immune  
 CC complex and an increase in mesangial cells. The animal model is useful  
 CC for analysing the pathology of chronic glomerular diseases and for  
 CC screening compositions for prevention and treatment of the diseases.  
 CC Highly uniform models can be made easily and in large numbers using  
 CC this method.  
 XX  
 SQ Sequence 1147 BP; 351 A; 239 C; 250 G; 307 T; 0 other;

Query Match 65.58; Score 748.6; DB 22; Length 1147;  
 Best Local Similarity 79.8%; Pred. No. 3.2e-215;  
 Matches 883; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

QY 37 TTCACTGTTTACAGAGATGATGACATCAAGGAAATGGAATGTGTTCTTTCTCT 96  
 DB 1 TTGACCTATTTCAGAGATGATGATAGACCAAGGAAATGGAATGTGTTCTTTCTCTCC 60  
 QY 97 CTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156

DB 61 CTGAGCATTTCTACTGCCCTGACCCCTAATCGCTCGGTGGTGTCTGAGGTGACTGTGCAGT 120  
 QY 157 CAGATTGATAAGTGTGCTTCATGTTAAACACTGCCCTCAGGATATGGAACTCTTCTTAATAGT 216  
 DB 121 CAGATTGACAGGCACTGCTGCTTTAAATACCATCAAGACAGGAACTCATCTTAATAAT 180  
 QY 217 CAGTCAGGGCTCCAGTCTCAACTGAAAGAGATTTTCTGATATAAATGATGATCCACAG 276  
 DB 181 CAGCCAGGACTTCAGTCAATGAAAGAGATTTCTGCTGATATAAATCACTCATCTCATAG 240  
 QY 277 GATTATGATCTCAGCATTTGTGAATGGGCTTTTCTGCTGAAAAGTGTATGCTTTTCATAG 336  
 DB 241 GATTATGATCTCAGCATTTGCCACTGGAGTTTTCAGGAAAGTCTATGATTTTCATAG 300  
 QY 337 GACTACATTGAGTGTGCCGAAAATTTATAGATGCCAAAGTGGAGCGAGTTGATTTAG 396  
 DB 301 AACTACATTGAGTGTGCTGAAAACCTTATCAATGCTTAAAGTGGAAAGTGGATTTTCA 360  
 QY 397 AATCAATTTAGAAAGCACTAGACGTAAATTAATAAGTGGTTGAAATGAAACACATGGC 456  
 DB 361 AATGATGTACAGATACAGATTTAAATTAATAAGTGGATTTGAAATGAGACACATGGA 420  
 QY 457 AAAATCAAGAACTGATTTGCTGAGGTGGCATAAGCTCATCTGCTGTAATGGTGTGTTG 516  
 DB 421 AAGATCAAGAGAGTGTGGGCGCAGCAGCTCAGCTCGTGGCTGTCTGAGTGTGTTG 480  
 QY 517 AATGCTGTGATCTTCAAGGCAAGTGGCAATCAGCTTCCCAAGAGGAAACATTAAT 576  
 DB 481 AAGCTGTGTTTACTTCAAGGCAATGGAATCGGCCCTTCCCAAGAGCTGATACCTCAGT 540  
 QY 577 TGCCATTTCAAATCTCCCAAGTGTCTTGGGAAGGAGTGGCCATGATGATCAGGAAGG 636  
 DB 541 TGCCGTTTGGTCTCCACGCTCTGGAAAAGTAGTTAATAAGTGGATCAAGAACGG 600  
 QY 637 AAGTCAATTTCTGTATTGAGGACCCATCAATCAAGATTTGAGCTTCAGATACAT 696  
 DB 601 CGGTTCAATTTGCTTACCATTGAGCAGCCCAATCAGGTTCTTGAGCTCCCAATATCAT 660  
 QY 697 GGTGGCAATAACATGTACGTTCTGCTGCTGAGAAATGACCTCTCTGAAATGAAACAA 756  
 DB 661 GGTGGCAATAACATGTACATTTATGCTGCTGAGGATGGCTTATGTGAAATGAAAGCAAG 720  
 QY 757 CTGACCTTTTCAAGATCTAATGGAATGGACCAATCCAAAGGCAATGACCTCTAAGTATGTT 816  
 DB 721 CTGAGTTTCCAGATCTGATGAGCTGGACCAATGAGGAGAAATGAAATCTCAGTATGTG 780  
 QY 817 GAGGTATTTTCTCTCAGTTCAAGATAGAGAGAAATTTGAAATGAAACAAATATTGAGA 876  
 DB 781 AAGCTGTTTCTCCCCCAGTTCAAGATAGAGAGAAATTTGAAATGACCCCACTTTGAAA 840  
 QY 877 GCCTTAGGCTGAAAGATCTTTGATGAATCCAAAGCAGATCTCTGGGATTTGCTTCG 936  
 DB 841 TCCTTAGGCTTGAAGATATCTTTGATGATGCTCAGTGGAGATCTCTTGGAATGCTCT 900  
 QY 937 GGGGTCGTCTGTATATATCAAGATGATGACAAATCTTACATAGAGGTCTCAGGAG 996  
 DB 901 GGAGTGTGCTCTCAGTATCAAGCTATGACAGTCAATTCATAGAGTCTCAGAGGAG 960  
 QY 997 GGCACGAGGCTACTGCTGCCACAGAGATTAATTTGTGAAAAGCAACTCCCTCTAGTCC 1056  
 DB 961 GGCACGAGGCTACTGCTGCCACAGAGAAATTAACATTTGTTGAAAAGCAGCTTCTCTAGTCC 1020  
 QY 1057 ACCTGTTTATAGAGCTGACACCCATCTCTATTTGTTATCAGGAGGATGACATCATCTTA 1116  
 DB 1021 ACAGTGTTCAGAGCCGACGCGCCCTTTCTGTTGTGATCAAGAGATGACATCATCTTA 1080  
 QY 1117 TTCACTGGCAAGTCTCTTGCCCTTGA 1143  
 DB 1081 TTTACTGGCAAGTCTCTTGCCCTTGA 1107

RESULT 13  
 ACA56226



ID ACAS6226 standard; cDNA; 316 BP.  
AC ACAS6226;  
XX  
DT 06-JUN-2003 (first entry)  
XX  
DE Human signalling pathway polynucleotide probe SEQ ID NO 824.  
XX  
XX Human; probe; ss; array element; Parkinson's disease;  
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;  
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.  
XX  
OS Homo sapiens.  
XX  
XX US6500938-B1.  
XX  
XX 31-DEC-2002.  
XX  
XX 30-JAN-1998; 98US-0016434.  
XX  
XX 30-JAN-1998; 98US-0016434.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Au-Young J, Seilhamer JJ;  
PI  
XX  
XX WPI; 2003-352189/33.  
XX  
XX Combination of polynucleotide probes, useful as array elements in a  
PT microarray for monitoring the expression of a number of target  
PT polynucleotides -  
XX  
XX Claim 1; SEQ ID NO 824; 65pp; English.  
XX  
XX The invention relates to a combination which, comprises a number of  
CC polynucleotide probes comprising a sequence selected from one of the 1490  
CC sequences mentioned in the specification. The combination is useful as an  
CC array element in a microarray for monitoring the expression of a number  
CC of target polynucleotides. The microarray is particularly useful in the  
CC diagnosis and treatment of cancer and immunopathology and neuropathology.  
CC The microarray is useful in diagnostics and treatment regimens, drug  
CC discovery and development, toxicological and carcinogenicity studies,  
CC forensics and pharmacogenomics. The microarray is also useful for  
CC monitoring progression of diseases and for developing sophisticated  
CC profiles for the effects of currently available therapeutic drugs. The  
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs  
CC and genomic fragments and in research and diagnostic applications. The  
CC array can detect changes in expression in a large number of genes coding  
CC for different signaling pathway populations which can be used to diagnose  
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,  
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease  
CC and Parkinson's disease. The present sequence represents a polynucleotide  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.  
XX  
XX Sequence 316 BP; 92 A; 63 C; 72 G; 88 T; 1 other;  
SQ

Query Match 26.6%; Score 304; DB 25; Length 316;  
Best Local Similarity 99.4%; Pred. No. 2.6e-81;  
Matches 315; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 822 ATTATTTCTCAGTTCAAGATAGAGAAGATTATGAATGAATGAATGAATTTGGAGCCCT 881  
DB 1 ATTATTTCTCAGTTCAAGATAGAGAAGATTATGAATGAATGAATTTGGAGCCCT 60  
QY 982 AGGCTGGAAGATATCTTTGATGAATCCAAAGCAGATCTCTCGGATTTCTTCGGGGG 941  
DB 61 AGGCTGGAAGATATCTTTGATGAATCCAAAGCAGATCTCTCGGATTTCTTCGGGGG 120  
QY 942 TCGTCTGTATATCAAGATGATCCAAATTTTACATPAGAGTCTCTGAGGGGAC 1001  
|||||

DB 121 TCGTCTGTATATCAAGATGATCCAAATTTTACATAGAGTCACTGAGGGGAC 180  
QY 1002 CGAGGCTACTGTGCTGCCACAGGAAGTAATTTCTAGAAAAGCAAACTCCCTCAGTCCAGGCT 1061  
DB 181 CGAGGCTACTGTGCTGCCACAGGAAGTAATTTCTAGAAAAGCAAACTCCCTCAGTCCAGGCT 240  
QY 1062 GTTTAGAGCTGACACCCATCTCTATTTCTATCAGGAAGATGACATCACTTATTTCAG 1121  
DB 241 GTTTAGAGCTGACACCCATCTCTATTTCTATCAGGAAGG-TGACATCATCTTATTTCAG 299  
QY 1122 TGGCAAAAGTTTCTGCC 1138  
DB 300 TGGCAAGTTTCTGCC 316

RESULT 14  
ACAS6060  
ID ACAS6060 standard; cDNA; 254 BP.  
XX  
AC ACAS6060;  
XX  
DT 06-JUN-2003 (first entry)  
XX  
DE Human signalling pathway polynucleotide probe SEQ ID NO 658.  
XX  
XX Human; probe; ss; array element; Parkinson's disease;  
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;  
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.  
XX  
OS Homo sapiens.  
XX  
XX US6500938-B1.  
XX  
XX 31-DEC-2002.  
XX  
XX 30-JAN-1998; 98US-0016434.  
XX  
XX 30-JAN-1998; 98US-0016434.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Au-Young J, Seilhamer JJ;  
PI  
XX  
XX WPI; 2003-352189/33.  
XX  
XX Combination of polynucleotide probes, useful as array elements in a  
PT microarray for monitoring the expression of a number of target  
PT polynucleotides -  
XX  
XX Claim 1; SEQ ID NO 658; 65pp; English.  
XX  
XX The invention relates to a combination which, comprises a number of  
CC polynucleotide probes comprising a sequence selected from one of the 1490  
CC sequences mentioned in the specification. The combination is useful as an  
CC array element in a microarray for monitoring the expression of a number  
CC of target polynucleotides. The microarray is particularly useful in the  
CC diagnosis and treatment of cancer and immunopathology and neuropathology.  
CC The microarray is useful in diagnostics and treatment regimens, drug  
CC discovery and development, toxicological and carcinogenicity studies,  
CC forensics and pharmacogenomics. The microarray is also useful for  
CC monitoring progression of diseases and for developing sophisticated  
CC profiles for the effects of currently available therapeutic drugs. The  
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs  
CC and genomic fragments and in research and diagnostic applications. The  
CC array can detect changes in expression in a large number of genes coding  
CC for different signaling pathway populations which can be used to diagnose  
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,  
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease  
CC and Parkinson's disease. The present sequence represents a polynucleotide  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.



```
Search completed: December 14, 2003, 01:19:47
Job time : 384 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 22:10:21 : Search time 89 Seconds  
(without alignments)  
5668.550 Million cell updates/sec

Title: US-09-508-997A-1  
Perfect score: 1143  
Sequence: 1 atggcctccctgtgcagc.....gcaactttctgccttga 1143

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1143	100.0	1950	2	US-08-472-659-30
2	1143	100.0	1950	2	US-08-474-661-30
3	1143	100.0	1950	2	US-08-611-977-30
4	304	26.6	316	4	US-09-016-434-824
5	254	22.2	254	4	US-09-016-434-658
6	225.6	19.7	1512	1	US-07-768-286B-3
7	222	19.4	1482	1	US-07-768-286B-5
8	220.6	19.3	1308	4	US-09-123-912-109
9	220.6	19.3	1308	4	US-09-643-597-109
10	220.6	19.3	1308	4	US-09-480-884A-109
11	220.6	19.3	1308	4	US-09-542-615A-109
12	220.6	19.3	1308	4	US-09-606-421B-109
13	211.6	18.5	1711	1	US-08-568-147B-1
14	210	18.4	1245	4	US-09-266-910-1
15	210	18.4	1245	4	US-09-266-910-2
16	206.8	18.1	1328	1	US-07-679-052A-14
17	206.8	18.1	1340	1	US-07-679-052A-16
18	206.8	18.1	1610	1	US-07-768-286B-1
19	206.8	18.1	2409	1	US-07-911-531-18
20	206.8	18.1	2409	1	US-07-693-636A-18
21	205	17.9	1419	4	US-09-123-912-111
22	205	17.9	1419	4	US-09-643-597-111
23	205	17.9	1419	4	US-09-480-884A-111
24	205	17.9	1419	4	US-09-542-615A-111
25	205	17.9	1419	4	US-09-606-421B-111
26	188.4	16.5	1278	4	US-10-024-427-1
27	188.4	16.5	1519	4	US-10-024-427-3

28	186.4	16.3	1366	3	US-09-200-965-1	Sequence 1, Appli
29	185.6	16.2	1361	4	US-09-016-434-1284	Sequence 1284, Ap
30	182.4	16.0	1230	4	US-09-613-303-54	Sequence 54, Appl
31	182.4	16.0	1873	4	US-08-545-573A-38	Sequence 38, Appl
32	182.4	16.0	2847	4	US-09-613-303-20	Sequence 20, Appl
33	179.6	15.7	1425	1	US-08-464-148-1	Sequence 1, Appli
34	179.6	15.7	1425	1	US-08-385-500-1	Sequence 1, Appli
35	179.6	15.7	1425	1	US-08-846-784-1	Sequence 1, Appli
36	170	14.9	1393	1	US-09-016-434-1075	Sequence 1075, Ap
37	170	14.9	1393	1	US-08-464-148-3	Sequence 3, Appli
38	170	14.9	1393	1	US-08-385-500-3	Sequence 3, Appli
39	170	14.9	1393	1	US-08-846-784-3	Sequence 3, Appli
40	156.8	13.7	2584	1	US-08-121-714-1	Sequence 1, Appli
41	156.8	13.7	2584	1	US-08-322-742-13	Sequence 13, Appl
42	156.8	13.7	2584	1	US-08-477-108A-1	Sequence 1, Appli
43	156.8	13.7	2584	1	US-08-477-112-1	Sequence 1, Appli
44	156.8	13.7	2584	5	PCT-US93-08322-1	Sequence 1, Appli
45	155.6	13.6	1152	1	US-08-315-831A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1  
US-08-472-659-30  
: Sequence 30, Application US/08472659  
: Patent No. 5831030  
: GENERAL INFORMATION:  
: APPLICANT: TSUJIMOTO, Masafumi  
: APPLICANT: IWASA, Fuyuki  
: APPLICANT: TSUBOUKKA, No. 5831030uo  
: APPLICANT: NAKAZATO, Hiroshi  
: APPLICANT: MIURA, Kenju  
: APPLICANT: ISHIDA, No. 5831030uhito  
: APPLICANT: KURIHARA, Tatsuya  
: APPLICANT: YAMAUCHI, Koza  
: APPLICANT: YAMAUCHI, No. 5831030omi  
: TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
: NUMBER OF SEQUENCES: 34  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Burns, Doane, Swecker & Mathis  
: STREET: P.O. Box 1404  
: CITY: Alexandria  
: STATE: Virginia  
: COUNTRY: United States  
: ZIP: 22313-1404  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent in Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/472,659  
: FILING DATE: 07-JUN-1995  
: CLASSIFICATION: 536  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: JP 4-212305  
: FILING DATE: 17-JUL-1992  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: JP 5-067339  
: FILING DATE: 04-MAR-1993  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/091,028  
: FILING DATE: 14-JUL-1993  
: ATTORNEY/AGENT INFORMATION:  
: NAME: McGowan, Malcolm K.  
: REGISTRATION NUMBER: 39,300  
: REFERENCE/DOCKET NUMBER: 001560-248  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 836-6620  
: TELEFAX: (703) 836-2021  
: INFORMATION FOR SEQ ID NO: 30:  
: SEQUENCE CHARACTERISTICS:

```

; LENGTH: 1950 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; STRAIN: A431
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "DNA coding for human
; OTHER INFORMATION: megakaryocyte differentiation factor."
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 74..1217
; US-08-472-659-30

```

```

Query Match      100.0%; Score 1143; DB 2; Length 1950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGCGCTCCCTGCTGCGCAAAATGCGAGATTTTGTCTCACTCTTCAGAGAGATGAT 60
Db      74  AUGGCTCCCTGCTGCGCAAAATGCGAGATTTTGTCTCACTCTTCAGAGAGATGAT 133

QY      61  GACATCAAGGAATGGAATGTGTCTTCTCTCTGAGCTCTTGGCTGCGCCCTGGCC 120
Db      134  GACAATCAAGGAATGGAATGTGTCTTCTCTCTGAGCTCTTGGCTGCGCCCTGGCC 193

QY      121  CTGCTCCGCTGGGCGCTCAAGTACTCCCTCTCTCAGATTTGATAGTTCCTCATGTT 180
Db      194  CTGCTCCGCTGGGCGCTCAAGTACTCCCTCTCTCAGATTTGATAGTTCCTCATGTT 253

QY      181  AACATGCTCAGGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240
Db      254  AACATGCTCAGGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 313

QY      241  AAAGAGTTTTTCTGTATATAATGATCCCAAGGATATGATCTCAGCATGTGAAAT 300
Db      314  AAAGAGTTTTTCTGTATATAATGATCCCAAGGATATGATCTCAGCATGTGAAAT 373

QY      301  GGGCTTTTGTGAAAAGTGTATGCTTCTCATTAAGGACTACATTTGATGCTGCGAATA 360
Db      374  GGGCTTTTGTGAAAAGTGTATGCTTCTCATTAAGGACTACATTTGATGCTGCGAATA 433

QY      361  TTATACGATGCCAAAGTGGAGCGAGTGTACTTTACGAATCATTTAGAGACACTAGACT 420
Db      434  TTATACGATGCCAAAGTGGAGCGAGTGTACTTTACGAATCATTTAGAGACACTAGACT 493

QY      421  AATATTAAATAGTGGTTGAAATGAAACATCGCAATCGGAAATCAAGACGTATGTGAA 480
Db      494  AATATTAAATAGTGGTTGAAATGAAACATCGCAATCGGAAATCAAGACGTATGTGAA 553

QY      481  GGTGGCATTAAGCTCATCTGCTGTAATGCTGTGTAATGCTGTACTTCAAGGCAAG 540
Db      554  GGTGGCATTAAGCTCATCTGCTGTAATGCTGTGTAATGCTGTACTTCAAGGCAAG 613

QY      541  TGGCAATACGCTTCCCAAGACGAAACCAATAATTCGCAATTCCTCCCAAGTGC 600
Db      614  TGGCAATACGCTTCCCAAGACGAAACCAATAATTCGCAATTCCTCCCAAGTGC 673

QY      601  TCTGGGAAGCGAGTGGCATGATGCTCAGGAACGGAAGTCAATTTCTGTATTGAG 660
Db      674  TCTGGGAAGCGAGTGGCATGATGCTCAGGAACGGAAGTCAATTTCTGTATTGAG 733

QY      661  GACCCATCAATGAATTTCTGAGCTTCAGATACATGCTGGCAATAAATGCTGCTCG 720
Db      734  GACCCATCAATGAATTTCTGAGCTTCAGATACATGCTGGCAATAAATGCTGCTCG 793

QY      721  CTGCTGAGATGACCTCTCGAAATTTGAAACAAACTGACCTTTTCAAGATCTAATGAA 780
Db      794  CTGCTGAGATGACCTCTCGAAATTTGAAACAAACTGACCTTTTCAAGATCTAATGAA 853

```

```

QY      791  TGGACCAATCCAGGCGAATGACCTCTTAAGTATCTTGAGTATTTTTTCTCCTCAAG 840
Db      854  TGGACCAATCCAGGCGAATGACCTCTTAAGTATCTTGAGTATTTTTTCTCCTCAAG 913

QY      841  ATAGAGAGCAATATGAAATGAAACAAATATTTGAGAGCTCTAGGCTCAAGATATCTTT 900
Db      914  ATAGAGAGCAATATGAAATGAAACAAATATTTGAGAGCTCTAGGCTCAAGATATCTTT 973

QY      901  GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTGGGGGGTCTGCTCTATATATCAAG 960
Db      974  GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTGGGGGGTCTGCTCTATATATCAAG 1033

QY      961  ATGATGCAAAATCTTACATAGAGTCTACTGAGAGGCGACCGAGGCTACTGTGCGACA 1020
Db      1034  ATGATGCAAAATCTTACATAGAGTCTACTGAGAGGCGACCGAGGCTACTGTGCGACA 1093

QY      1021  GGAAGTAATATTTAGAAAGCAACTCCCTCAGTCCACGCTTTTAGAGCTGACCAACCA 1080
Db      1094  GGAAGTAATATTTAGAAAGCAACTCCCTCAGTCCACGCTTTTAGAGCTGACCAACCA 1153

QY      1081  TTCTTATTTTATCAGAAAGGATGACATCATCTTATTCAGTGGCAAAAGTTCTTGGCCT 1140
Db      1154  TTCTTATTTTATCAGAAAGGATGACATCATCTTATTCAGTGGCAAAAGTTCTTGGCCT 1213

QY      1141  TGA 1143
Db      1214  TGA 1216

```

## RESULT 2

```

US-08-474-661-30
; Sequence 30, Application US/08474661
; Patent No. 5874253
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. 5874253uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5874253uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5874253omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,661
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRICE APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: REA, TERESA STANEK
; REGISTRATION NUMBER: 30,427

```





```

/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0002 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 658:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 254 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KERANOT02
/ CLONE: 2622817
/ US-09-016-434-658

Query Match          22.2%; Score 254; DB 4; Length 254;
Best Local Similarity 100.0%; Pred. No. 7.7e-73;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      303   GCCTTTTGGTGAAAAGGTGATGGCTTTCAATAGACACTACATGAGTGTGCCGAAAAATT    362
Db      1     GCCTTTTGGTGAAAAGGTGATGGCTTTCAATAGACACTACATGAGTGTGCCGAAAAATT    60

QY      363   ATACGATGCCAAGTGGAGCGAGTTCACITTTACGAATCATTTAGAAGACACTAGACGTAA    422
Db      61   ATACGATGCCAAGTGGAGCGAGTTCACITTTACGAATCATTTAGAAGACACTAGACGTAA    120

QY      423   TATTATAAAGTGGGTTGAAAATGAACAACATGGCAGAAATCAAGAACGTGATTTGGTGAAGG    482
Db      121   TATTATAAAGTGGGTTGAAAATGAACAACATGGCAGAAATCAAGAACGTGATTTGGTGAAGG    180

QY      483   TGCGATAGACTCATCTGCCTGATGCTGTGATGCTGTGCTGTGNATGCTGTACTTCGAAGGCAAGT    542
Db      181   TGCGATAGACTCATCTGCCTGATGCTGTGATGCTGTGCTGTGNATGCTGTACTTCGAAGGCAAGT    240

QY      543   GCAATCAGCCTTCA 556
Db      241   GCAATCAGCCTTCA 254

RESULT 6
US-07-768-286B-3
/ Sequence 3, Application US/07768286B
/ Patent No. 5444153
/ GENERAL INFORMATION:
/ APPLICANT: GOSS, Neil H.
/ APPLICANT: RICHARDSON, Michael A.
/ TITLE OF INVENTION: VARIANTS OF PAI-2
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ APPLICATION NUMBER: US/07768,286B
/ FILING DATE: 19911011
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:

```



APPLICATION NUMBER: PCT/AU90/00603  
 FILING DATE: 20-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 16786/157 CHAC  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1512 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHEetical: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 CELL TYPE: Monocyte  
 CELL LINE: U937  
 IMMEDIATE SOURCE:  
 CLONE: BTA 1916  
 FEATURE:

NAME/KEY: CDS  
 LOCATION: 22..1200  
 IDENTIFICATION METHOD: experimental  
 OTHER INFORMATION: /function= "Product binds to  
 OTHER INFORMATION: urokinase, tissue plasminogen activator"  
 OTHER INFORMATION: /product= "PAI-2 variant, protease sensitive site  
 OTHER INFORMATION: removed"  
 OTHER INFORMATION: /evidence= EXPERIMENTAL  
 OTHER INFORMATION: /notes= "Codes for human plasminogen activator  
 OTHER INFORMATION: inhibitor type 2 protein in which amino acids 74  
 OTHER INFORMATION: to 96 inclusive have been deleted."  
 US-07-768-286B-3

Query Match 19.7%; Score 225.6; DB 1; Length 1512;

Best Local Similarity 51.6%; Pred. No. 4.5e-63;

Matches 590; Conservative 0; Mismatches 524; Indels 30; Gaps 2;

QY 1 ATGGGCTCCCTGCTGAGCAATGCAAGATTTGCTTCAACCTGTTCAAGAGATGAT 60  
 DB 22 ATGGAGATCTTTGTGGCAACACACTTTTGGCCCTCAATTTATTCAGCATCTGGCA 81  
 QY 61 GACAAATCAAGGAATGAATGTTCTTTCTCTGAGCCTCTTCTGCTGCTGCTGCTG 120  
 DB 82 AAAGCAAGCCCAACAGACCTCTTCTCTCCCATGGAGCATCTGCTCCACATGGCC 141  
 QY 121 CTGGTCGGCTGGGGCTCAAGATGATGATGCTCTCTCAGATGATGATGATGATGAT 180  
 DB 142 ATGGTCTACATGGGCTCCAGGGGCGAGCACCAAGAGCAGATGGCCAGGTGCTTCAGTT 201  
 QY 181 AACACTGCTCAGATATGAACTCTTTCTAATAGTCAGTCAGGGCTCCAGTCTCACTG 240  
 DB 202 ATGAAGTGGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 261  
 QY 241 AAAAGAGTTT-----TCTGATATAAATGATCCCAAGAGGATTTATGATCTC 289  
 DB 262 CATTATCTTCCGCTCTCTCAGCTCTGCAATCAATGATCCAGAGGATTTATGATG 321  
 QY 289 AGCATTTGAATGGGCTTTTGTGAAAAGTGTATGGTTTCTATAGGCTCATATGAG 348  
 DB 322 GAAAGTGTCAATAGCTTTTGTGAGAAGTCTGCGAGTTCGCGAAGAATATATCGA 381  
 QY 349 TGTGCGAATATATAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 408  
 DB 382 CTCTGTCAGAAATATTACTCTCAGACCCCGAGGAGTAGACTTCTTAGAATGTGAGAA 441  
 QY 409 GACACTPAGACGTAATATTAATAGTGGGTTGAAAATGAACACATGGCAAAATCAAGAC 468

DB 442 GAAGCTAGAAAAGATTAATTTCTTGGGTCAAGACTCAACCAAGGCAAAATCCCAAC 501  
 QY 469 GTGATTGGTGAAGTGGCAATAGCTCACTGTGTAAATGGTGTGGTGAATGCTGTGAC 528  
 DB 502 TTGTTACCTGAAGTTCCTAGATGGGATACCAAGATGGTCTCTGGTGAATGCTGTCTAC 561  
 QY 529 TTCAAGGCAAGTGGCAATCAGCCTTCCCAAGAGGCAACCAATTAATTCGCATTTCAA 588  
 DB 562 TTCAAGGAAAGTGGAAAATCTCCAATTGAGAGAAACTAAATGGGCTTTATCTTTCCGT 621  
 QY 589 TCTCCCAAGTGTCTCTGGGAGGAGTGCATGATGATGATGATGATGATGATGATGAT 648  
 DB 622 GTAAACTGGCTCAGCGCACCTGTACAGATGATGATGATGATGATGATGATGATGAT 681  
 QY 649 TCTGTTATTGAGGACCAATCAATGAATTTTGAAGCTTCAAGTCAAGTCAAGTCAAGT 708  
 DB 682 GGATACATAGAGACCTTAAGGCTCAGATTTAGAACTCCCATATGCTGAGATGTTAGC 741  
 QY 709 ATGTACGTTCTGCTGCTGA-----GAATGACCTCTCTGAATTTGAA 750  
 DB 742 ATGTTCTTGTGCTTCCAGATGAATTTCCGATGTGTCCATGCTGGCTTGGAGCTGTGGA 801  
 QY 751 AACAACTGACCTTTCAAGATCTAATGAATGGACCAATCCAGGCGGATGACCTTAAG 810  
 DB 802 AGTGAATATACCTATGACAACTCTCAAGTGGACCCAGCAAGACAAAATGGCTGAAGAT 861  
 QY 811 TATGTTGAGGTATTTTCTCAGTTCAAGATGAGAGAAATTAATGAATGAACAAATAT 870  
 DB 862 GAAGTTGAGGTATACATACCCAGTTCAAAATTAGAGAGCATTATGAATCAGATCCATT 921  
 QY 871 TTGAGAGCCCTAGGCTGGAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGAT 930  
 DB 922 CTGAGAGCATGGGCAATGGAGGAGCCCTTCAAGAGGAGGAGGAGGAGGAGGAGGAG 981  
 QY 931 GCTTCGGGGGCTGCTGTGTATATATCAAGAGATGATGATGATGATGATGATGATGAT 990  
 DB 982 TCGAGAGGAATGACCTGTTTCTTCTGAGTGTTCACCAAGCCATGCTGGATGTGAAT 1041  
 QY 991 GAGAGGGCAGGAGGCTACTGCTCCACAGAGAGTAAATTTGTAGAAAGCAACTCCCT 1050  
 DB 1042 GAGAGGGCACTGAAGCAGCCGCTGGCAGAGAGGTGTTATGACAGGAGAACTGGACAT 1101  
 QY 1051 CAGTCCAGCTGTTTAGAGTGCACCCATCTCTCTTTGTTATCAGGAAGGATGATC 1110  
 DB 1102 GAGGCCACAGATTGTGGAGATCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1161  
 QY 1111 ATCT 1114  
 DB 1162 AACT 1165

## RESULT 7

US-07-768-286B-5  
 ; Sequence 5, Application US/07768286B  
 ; Patent No. 544153  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GOSS, Neil H.  
 ; APPLICANT: RICHARDSON, Michael A.  
 ; TITLE OF INVENTION: VARIANTS OF PAI-2  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/768,286B  
 FILING DATE: 19911011  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/AU90/00603  
 FILING DATE: 20-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 16786/157 CHNC  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1482 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 CELL TYPE: Monocyte  
 CELL LINE: U937  
 IMMEDIATE SOURCE:  
 CLONE: BTA 1922  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 22..1170  
 IDENTIFICATION METHOD: experimental  
 OTHER INFORMATION: /function= "Product binds to  
 OTHER INFORMATION: urokinase, tissue plasminogen activator"  
 OTHER INFORMATION: /product= "PAI-2 variant, protease sensitive site  
 OTHER INFORMATION: removed"  
 OTHER INFORMATION: /evidence= EXPERIMENTAL  
 OTHER INFORMATION: /note= "Codes for human plasminogen activator  
 OTHER INFORMATION: inhibitor type 2 protein in which amino acids 66  
 OTHER INFORMATION: to 98 inclusive have been deleted."  
 US-07-768-286B-5

Query Match 19.4%; Score 222; DB 1; Length 1482;  
 Best Local Similarity 52.2%; Pred. No. 6,5e-62;  
 Matches 591; Conservative 0; Mismatches 505; Indels 36; Gaps 3;

QY 1 ATGGCTCCCTGCTGCGAGCAATGCGAGTTTGGCTCAACCTCTTCAGAGATGGAT 60  
 DB 22 ATGGAGATCTTTGTGGCAACACACTCTTGGCCCTCAATTTATTCAAGCATCTGGCA 81  
 QY 61 GACAATCAAGGAATGGAATGTGTCTCTCTCTGAGCTCTTCGGTCCCTGGCC 120  
 DB 82 AAAGCAGCCCCCAGACACTCTCTCTCCCATGGAGCATCTGTCACCATGGCC 141  
 QY 121 CTGGTCCGCTGGGGCTCAAGATGACATCCCTCTCTCAAGATTGATAGTTGCTCATGTT 180  
 DB 142 ATGGTCTACATGGGCTCCAGGGGCGACACCGAAGACAGATGCCAAGGCTTCAGTTT 201  
 QY 181 AACATGCTCAGGATATGAACCTCTCTTAATAGTCAGTCAGGCTCCAGTCTCACTG 240  
 DB 202 ATGAAGTGGAGCGCTGAGCA-----TAAATCATCTCATCTCCGCTCTC----- 250  
 QY 241 AAAAGAGTTTTTCTGATATAATGATCCCAAGGATATGATCTCAGCATTTGGAAT 300  
 DB 251 -----TCAGCTCTGCAATCAATGATCCACAGGAATTTTACTGGAAGATGCAAT 303  
 QY 301 GGCTTTTTCGAAAAAGGTATAGGCTTTTCATAGGACTACATTTAGTGTGCGGAAAAA 360  
 DB 304 AAGCTTTTGGTGAAGTCTCGAGCTTCGGGAAGATATATTCGACTCTCTCAGAA 363  
 QY 361 TTATAGATGCAAGTGGAGCGATTCATTTACGATCATTTAGAGACATAGAGCT 420

DB 364 TATTACTCTCTCAGAACCCCGCAGTAGACTTCTTAGAATGTCGACAGAACTAGAAA 423  
 QY 421 ATATTATTAAGTGGTGGTGAATAATGAAAACATGCGCAAAATCAAGAACGTGATTGGTGA 480  
 DB 424 AAGATTAAATCTCTGGTCAAGACTCAAAACAAAGGCAAAATCCCAAACTTGTACCTGAA 483  
 QY 481 GGTGGCATAGCTCATCTGCTGTAATGCTGCTGGTGAATGCTGTGTACTTCAAGGCAAG 540  
 DB 484 GTTCTGTAGATGGGATACCAAGATGGTCTGCTGGTGAATGCTGTACTTCAAGGAAAG 543  
 QY 541 TGGCAATCAGCTTCCACCAAGAGCAAAACATAAATTTCCATTTTCAAAATCTCCCAATGC 600  
 DB 544 TGGAAATCTCAATTTGAGAAGAACTAAATGGGCTTTATCCTTTCCGTGTAACCTCGCT 603  
 QY 601 TCTGGAGGCGATCGCCATGATGATCATCAAGAGCAAGTTCATTTGCTGTATTGAG 660  
 DB 604 CAGCGCACCTGTACAGATGATGATCTTGGCTGAAAGCTAAACATTGGATACATAGAA 663  
 QY 661 GACCCATCAATGAAGATTTCTGAGCTCAGATACATATGCTGCATATAAATGCTAGCTTCTG 720  
 DB 664 GACCTAAGGCTCAGATTTCTAGAACTCCCATATGCTGGAGATGTAGCATGTTCTTGTG 723  
 QY 721 CTGCTCA-----GAATGACCTCTCTGAAATTCGAAATCAAAACAACTGACC 762  
 DB 724 CTTCCAGATGAAATTCGCGATGTGTCCACTGGCTTGGAGCTGCTGAAAGTGAATAACC 783  
 QY 763 TTTCCAGATCTAATGGAATGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTA 822  
 DB 784 TATGACAACTCAACAGTGGACCGCAAGACAAATGCTGAAGATGAAGTTGAGGTA 843  
 QY 823 TTTTCTCTCAGTTCAAGATAGAGAAATATGAAATGAAACAAATTTTGGAGCCCTA 882  
 DB 844 TACATACCCAGTTCAMATTAGAGAGCATATGACTCAGATCCATCTCAGAGCATG 903  
 QY 983 GGGCTGAAAGATATCTTTGATGATCCAAAGCAGATCTCTCGGATGCTTCGGGGGGT 942  
 DB 904 GGCATGGAGAGCGCTTCAAAAGGAGCGGCCAATTTCTCAGGATGTCGAGAGGAAAT 963  
 QY 943 CGTCTGATATATCAAGGATGATGACAAATCTTACATAGAGTCACTCAGAGGCGACC 1002  
 DB 964 GACCTGTTCTTTCTGAAGTGTCCACCAAGCCATGGTGAATGTGAATGAGAGGGGCACT 1023  
 QY 1003 GAGCTACTGCTGCCACAGGAAGTAATATTGTAGAAAGCAACTCCTCAGTCCACGCTG 1062  
 DB 1024 GAGCAGCGCTGGCAAGAGGTGTATGACAGGGAGAACTGGACATGGAGGCCACAG 1083  
 QY 1063 TTTAGAGTGCACCCATCTCTATTTGTTATCAGGAAGGATGACATCATCT 1114  
 DB 1084 TTTGGCAGATCATCTCTTTCTTTTCTTTTATTATGATAGATAACCAACT 1135

RESULT 8  
 US-09-123-912-109  
 ; Sequence 109, Application US/09123912A  
 ; Patent No. 6312695  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Wang, Tongcong  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER  
 ; FILE REFERENCE: 210121.455C1  
 ; CURRENT APPLICATION NUMBER: US/09/123,912A  
 ; CURRENT FILING DATE: 1998-07-27  
 ; PRIOR APPLICATION NUMBER: 09/040,802  
 ; PRIOR FILING DATE: 1998-03-18  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: Patent in Ver. 2.0  
 ; SEQ ID NO 109  
 ; LENGTH: 1308  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-123-912-109  
 Query Match 19.3%; Score 220.6; DB 4; Length 1308;



Db 442 GCAGCGAGTAAAGTGGAGAGAGATTAATTCCTGGGTGGAAGCAAAACAATGAAGA 501  
QY 460 ATCAAGACGATGATGTTGAAGTGGCAATAGCTCATCTGCTGTAATGTTGCTGGTGAAT 519  
Db 502 ATCAAGACGATGATGTTGCAAGATGCTCTATTAGTAGCTTACCAAGCTGGTGGTGAAC 561  
QY 520 GCTGTGTACTTCAAGGCAAGTGGCAATCAGCTTACCAAGAGCGAAACCAATGAATGC 579  
Db 562 ATGGTTTATTTAAAGGCAATGGGACAGGAGTTTAAAGAAATAATAAGGAAGAG 621  
QY 580 CATTTCAATCTCCCAAGTCTCTGGGAAGGCAAGTCCCATGATGATGATCAGGAAGGAG 639  
Db 622 AAATTTGGATGAATAAGACACAAAGTAAATCTGTACAGATGATGACACAGAGCCATTCC 681  
QY 640 TTCAATTTGCTGTATTTAGAGACCCATCAATGAAGATTTCTTGAGCTCAGAT---ACAAT 696  
Db 682 TTTAGCTTCACTTTCTCGAGGACTTGCAGGCCAAAATCTTAGGGATTCATATATAAAAC 741  
QY 697 GTGGCATAAAATGTTAGCTTCTGCTGCTGA-----GAATGACCTCTCTGAATGAA 750  
Db 742 AACGACCTAAGCATGTTTGTGCTTCTGCCAAACGATGATGCTGGAGAGATAATA 801  
QY 751 AACAACTGACCTTTCAGATCTTAATGAATGGAATGGAACCAATCCAAAGCGAATGACCTTAAG 810  
Db 802 GATAAAATAAGTCTCGAGAAATTTGGTAGGTGGACTAGTCCAGGCGCATATGGAAGAAGA 861  
QY 811 TATGTTGAGGTATTTTTCTCAGTCTAAGATAGAGAAATATGAATGAACAAATAT 870  
Db 862 AAGGTGAATCTGCACTTGGCCGGTTTGAGGTGGAGACAGTTACGATCTAGAGCGGCTC 921  
QY 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCGGAT 930  
Db 922 CTGGCTGCATGGGATGGCGATGCTTCAAGTGGACCAAAAGCCGCTACTCGGAATG 981  
QY 931 GCTTCGGGGTCTGTGTATATCAAGGATGATGACCAAAATCTTACATAGAGTCACT 990  
Db 982 TGTTCAGGCTCGGGTGTACGCCAGAGTTCTGCACAGTTCTTGTGGCAGTAACT 1041  
QY 991 GAGGAGGCGCAGGCTACTGCTGCCACAGAGTAATATGTTAGAAAGCACTCCCT 1050  
Db 1042 GAGGAGGCGCAGGCTGCTGCCACAGGCTGCTGCCACAGGCTTACTGTTCATCAGCACA 1101  
QY 1051 CAGTCCAGGCTGTTTAGAGCTGAACCCATCTCTATTTGTTATCAGAGGATG----- 1105  
Db 1102 GGTATGAATATGTTCACTCCATCATCCCTCTGTTCTTCAATCAGCACAATGAATCC 1161  
QY 1106 -ACATCATTTATTCAGTGGCAAAAGTTTCTTGGCCCTTGA 1143  
Db 1162 AACAGCATCTCTCTTCGGCAGATTTTCTTCTCCNTAA 1200

## RESULT 10

US-09-480-884A-109

; Sequence 109, Application US/0948084A

; Patent No. 6482597

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; FILE REFERENCE: 210121.45506

; CURRENT APPLICATION NUMBER: US/09/480.884A

; CURRENT FILING DATE: 2001-08-27

; NUMBER OF SEQ ID NOS: 330

; SOFTWARE: FastSeq for Windows Version: 3.0

; SEQ ID NO 109

; LENGTH: 1308

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-480-884A-109

Query Match 19.3%; Score 220.6; DB 4; Length 1308;

Best Local Similarity 52.7%; Pred. No. 1.7e-61;

Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

QY 1 ATGCGCTCCCTGCTGCGAGCAATGCAAGTGGTTCCTCAACTGTTCCAGCTCTTCAGAGAGTGGAT 60  
Db 25 ATGGATTCACCTTGGCGCGCTGAGCTCGACTTGGGTTTGATCTTTTCAAGAGCTGAAG 84  
QY 61 GACAAACAAGAAATGGAAATGTGTTTCTCTCTGAGCTCTCTCGTGCCTGGCC 120  
Db 85 AAACAAATG---ATGCGACATCTCTTTTCCCTCTGGGCACTTTCAGTGCATTTGGC 141  
QY 121 CTGTCGGTCTGGGGCTTAAGATGATCTCCCTCTCTCAGATTGAATGTTCTCATGTT 180  
Db 142 ATGTCCTCTCTGGGACCCGAGGAGCCACCTTCCAGTTGGAGGAGTGTTCCTCTCT 201  
QY 181 AACACTGCGCTCA-----GGATATGGAACCTCTCTTAATAGTCAG 219  
Db 202 GAATAAGAGACAGAGAGCTCAAGATPAAGGCTGAAGAAAGAGTGATTTGAGACACA 261  
QY 220 TCAGGCTCCAGTCTCAACTGAAAGAGTTCCTTCTGATATAATTCATCCCAAGAGAT 279  
Db 262 GAAGCAGTACATCAACAATTCGAAAGTTCCTGACTGAAATAGCAAACTCACTAATGAT 321  
QY 280 TATGATCTCAGCATTTGAATGGGCTTTTGTGTAAGAGTGTATGGCTTTTCATAGGAC 339  
Db 322 TATGAATCGAAATGAACCAACAGGCTGTTTGGAGAAACATACCTCTCTTCCATCAAAA 381  
QY 340 TACATTTAGTGTGGGAAATATATACGATGCCAAAGTGGAGGAGTTGACTTTACGAAT 399  
Db 382 TACTTATATATGTTGAAATATATATCATGATCTCTGSAACCTGTTGATTTGTAAT 441  
QY 400 CATTTAGAAGACATAGACGCTAATATTAATTAAGTGGCTTCAAAATGAACACATGSCAA 459  
Db 442 GCAGCGATGAAGTTCGAAAGAGATTAATTCCTGGGTTGAAAGCAAAACAATGAAGAA 501  
QY 460 ATCAAGACGATGATGTTGAAGTGGCAATAGCTCATCTCTGTAATGCTGTTGGTGAAT 519  
Db 502 ATCAAGGACTTGTTCACAGATGCTCTATTAGTAGCTTACCAAGCTGGTCTGTTGGTGAAC 561  
QY 520 GCTGTCTACTTCAAGCAAGTGGCAATCAGCTTACCAAGAGGCAAAACCAATATTCG 579  
Db 562 ATGTTTATTTTAAAGGCAATGGGACAGGAGTTTGAAGAAATAATAAGGAAGAG 621  
QY 580 CATTTCAATCTCCCAAGTCTCTGGGAAGGCTCCCATGATGATCAGTAAGCAAGAG 639  
Db 622 AAATTTGGATGAATAAGACACAAAGTAAATCTGTACAGATGATGACACAGAGCATTC 681  
QY 640 TTCAATTTGCTGTATTTGAGGACCCATCAATGAAGTTCCTGAGTCTCAGT---ACAAT 696  
Db 682 TTTAGCTTCACTTTCTCGAGGACTTGCAGGCCAAAATTCCTAGGATTCATATAAANAAC 741  
QY 697 GTGGCATAAAACATGATGCTGCTGCTGA-----GAATGACCTCTCTGAATGAA 750  
Db 742 AACGACCTAAGCATGTTTGTGCTTCTGCCAAACGATGATGCTGGAGAGATAATA 801  
QY 751 AACAACTGACCTTTCAGATCTTAATGAATGGAATGGAACCAATCCAAAGCGAATGACCTTAAG 810  
Db 802 GATAAAATAAGTCTCGAGAAATTTGGTAGGTGGACTAGTCCAGGCGCATATGGAAGAAGA 861  
QY 811 TATGTTGAGGTATTTTTCTCAGTCTAAGATAGAGAAATATGAATGAACAAATAT 870  
Db 862 AAGGTGAATCTGCACTTGGCCGGTTTGAGGTGGAGACAGTTACGATCTAGAGCGGCTC 921  
QY 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCGGAT 930  
Db 922 CTGGCTGCATGGGATGGCGATGCTTCAAGTGGACCAAAAGCCGCTACTCGGAATG 981  
QY 931 GCTTCGGGGTCTGTGTATATCAAGGATGATGACCAAAATCTTACATAGAGTCACT 990  
Db 982 TGTTCAGGCTCGGGTGTACGCCAGAGTTCTGCACAGTTCTTGTGGCAGTAACT 1041

```
QY 991 GAGAGGGGACCCAGGCTACTGCTGGCCAGAGAGTAAATATGTAGAAAGCAACTCCCT 1050
    |||||
Db 1042 GAGGAGGACCCAGGCTGAGCTGCCACTGGCAGTAGGCTTTTACTGTCATCCGCCCA 1101
    |||||
QY 1051 CAGTCCAGGCTGTTAGAGCTGACCCCAATTCCTATTGTTATCAGGAAGATG----- 1105
    |||||
Db 1102 GGTATGAAATGTTCACTGCAATCATCCCTTCCTGTTCTTCATCAGGCACATGAATCC 1161
    |||||
QY 1106 -ACATCATCTTATTAGTGGCAAAAGTTTCTTCCCTTGA 1143
    |||||
Db 1162 AACAGCATCTCTTCTTGGCAGATTTTCTTCCCTTAA 1200
    |||||

RESULT 11
US-09-542-615A-109
; Sequence 109, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy R.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542.615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 109
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-109

Query Match 19.3%; Score 220.6; DB 4; Length 1308;
Best Local Similarity 52.74; Pred. No. 1.7e-61;
Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

QY 1 ATGGCTCCCTTGTGCGAGCAATGCGAGATTTTGCTTCAACCTGTTCCAGAGATGGAT 60
Db 25 ATGGATTCACTTGGCCCGCTCAGACTCGACTTGGTTGCTTTTCAAGAGCTGAG 84
QY 61 GACATCAAGAAATGAAATGTTCTTTCCTCTCTGAGCTCTTCTGTCGCTCGCC 120
Db 85 AAACAAATG---ATGCAACATCTTTTTCCTCTGCTGGCATCTTCACTGCAATGGC 141
QY 121 CTGGTCCGCTTGGGCGCTCAAGATGACTCCCTCTCAGATGATAGTTGCTTCACTGT 180
Db 142 ATGGTCTCTTGGGACCCGAGAGCCACCGCTTCCCGATTGGAGAGGTGTTTCACTCT 201
QY 181 AACACTGGCTCA-----GGATATGAAACTCTTCTAATAGTAG 219
Db 202 GAAAGAGAGAGAGAGCTCAGAAATAAGCTGAAGAAAGAGGTGATTGAGAACACA 261
QY 220 TCAGGCTCCAGTCTCACTGAAAGAGTTTTCCTGATATATATGATCCCAAGAT 279
Db 262 GAAGCAGTACATCAACAAATTCGAAAGTTTTCGACTGAAATAAGCAAACTCAATGAT 321
QY 280 TATGATCAGATTTGATGGCTTTTTCGTAAGATGTTGCTTTCATAGGAC 339
Db 322 TATGAACTGACATTAACCAACAGGCTTTTGGAGAAATAACATACCTTCTCTTCAAAA 381
QY 340 TACATGAGTGGCGGAAAAATATACGATCCCAAGTGGAGCGAGTACTTTAGCAAT 399
Db 382 TACTTAGATTATGTAATAATATATCATGCACTCTGGAACCTGTGATTTTGAAT 441
QY 400 CATTTAGAGACACTAGACGTAAATATTAATAGTGGTTGAAATGAAACACATGCAAA 459
Db 442 GCAGCGGATGAAGATCGAAAGAGATTAACTCTGGGTTGAAGGCAAAACAAATGAAAA 501
    |||||
```

```
QY 460 ATCAAGAACGTATGTTGAGAGTGGCATTAAGCTCATCTGCTGTAAATGTTGTTGAT 519
    |||||
Db 502 ATCAAGAGCATTTGTTCCAGATGGCTCTATTAGTAGTCTTACCAAGCTGTTGTTGTAAC 561
    |||||
QY 520 GCTGTGTACTTCAAGGCAAGTGGCAATCAGCTTCCACCAAGAGCGGAAACCAATTAATGC 579
    |||||
Db 562 ATGGTTTATTTTAAAGGCAATGGGACAGGAGTTTAAAGAAATATCTAAGGAAGAG 621
    |||||
QY 580 CATTTTCAATCTCCCAAGTCTCTGGGAAGGCGAGTCGCGCATGATGCATCAGGAAGCGAAG 639
    |||||
Db 622 AAATTTGGATGAATAAGAGCACAAAGTAAATCTGTACAGATGATGACACAGAGCCATTC 681
    |||||
QY 640 TTCAATTTGCTGTATTGAGGACCATCAATGAAGATTTCTTGAGCTCAGAT---ACAAT 696
    |||||
Db 682 TTTAGCTTCACTTCTCTGGAGACTTTCAGGCCAAATTTCTAGGATTCATATTAATAAC 741
    |||||
QY 697 GTGGCATAAACATGTACCTTCTGCTGCTGA-----CAATGACCTCTCTGAAATGAA 750
    |||||
Db 742 AACGACCTAAGCATGTTTGTGCTTCTGCCAAGCAGCATGATGGCTGGAGAGATAATA 801
    |||||
QY 751 ACAAACCTGACCTTTCAGAACTAATGAATGGACCAATCCAGGCGAATGACCTCTAAG 810
    |||||
Db 802 GATAAATAAGTCTCTGAGAAATTTGAGAGTGGACTAGTCCAGGSCATATGGAAGAAAGA 861
    |||||
QY 811 TATGTTGAGGTATTTTCTCTCAGTTCAGATAGAGAAATTAATGAATGAACAAATAT 870
    |||||
Db 862 AAGTGAATCTGCACTTGGCCCGTTTTCAGGTGAGAGCAGTTAGCATCTAGAGCGGTC 921
    |||||
QY 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGCAGAGATCTCTCTGGATT 930
    |||||
Db 922 CTGGCTGCCATGGGATGGCGATGCCCTTCAGTGAGCACAAGCGCACTACTCGGAAATG 981
    |||||
QY 931 GCTTCGGGGGTCTCTGTATATATCAGGATGATGCAAAATCTTACATAGAGTCACT 990
    |||||
Db 982 TCGTCAGGCTCCGGGTTGTACGCCAGAGTTCTCTCAGCTTCTTCTCATCAGGCACATGAATCC 1041
    |||||
QY 991 GAGAGGGCAGCGAGGCTACTGCTCCACAGGAAGTAATATTGTAGAAAAGCACTCCCT 1050
    |||||
Db 1042 GAGGAAGGACCGAGGCTGCACTGCCACTGGCATAGGCTTTTACTGTCACTCCGCCCA 1101
    |||||
QY 1051 CAGTCCACGCTGTTTAGAGTGCACCACTTCTTCTATTTTATCAGGAAGGATG----- 1105
    |||||
Db 1102 GGTATGAAATGTTTCACTGCAATCATCCCTTCTGTTCTTCTCATCAGGCACATGAATCC 1161
    |||||
QY 1106 -ACATCACTTATTCACTGSCAAAGTTTCTTCCCTTGA 1143
    |||||
Db 1162 AACAGCATCTCTTCTTCCGAGATTTTCTTCTCTCTTAA 1200
    |||||

RESULT 12
US-09-606-421B-109
; Sequence 109, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606.421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 109
; LENGTH: 1308
```

TYPE: DNA  
ORGANISM: Homo sapien  
US-09-606-421B-109

Query Match 19.3%; Score 220.6; DB 4; Length 1308;  
Best Local Similarity 52.7%; Pred. No. 1.7e-61;  
Matches 521; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

QY 1 ATGGCTCCCTGCTGCGAGCAATGCAGAGTTTGGCTTCAACCTGTCAGAGAGTGGAT 60  
DB 25 ATGGATCTACTTGGGCGCTGACACTCGACTTGGGTTTGCATCTTTTCAAGAGCTGAAG 84  
QY 61 GACAATCAAGGAATGAAATGTGTTTCTCTCTGAGCCTCTTCGCTGCCCTGGCC 120  
DB 85 AAAACAATG--ATGCAACATCTCTTTCCCTGTGGCACTTGAATGCAATGGC 141  
QY 121 CTGGTCGGTGGCGCTCAAGAGCTCCCTCTCTCAAGTGAATGATAGTTGCTTCACT 180  
DB 142 ATGGTCTCTCTGGGACCCGAGAGCCACCGCTTCCAGTTGGAGAGTGTTCACCT 201  
QY 181 AACACTGGCTCA-----GGATNTGGAACCTCTCTAATAGTCAG 219  
DB 202 GAAAAGAGAGAGAGCTCAAGATTAAGCTGAGAAAAGAGGTGATGAGAACACA 261  
QY 220 TCAGGCTCCAGTCTCACTGAAAAGAGTTTCTGATATAATGATGCCACAGAGAT 279  
DB 262 GAAGCAGTACATCAACAATTCAAAAGTTTGTGACTGAAATTAAGCAACTCACTATGAT 321  
QY 280 TATGATCTCAGCATGTGAATGGCTTTTCTGMAAAGTGTATGGCTTTCATAGGAC 339  
DB 322 TATGAACCTGAACATAACCAACAGGCTGTTTGGAGAAAACATACCTCTCTCAAAA 381  
QY 340 TACATGTAGTGTGCGAAAAATTTATACGATGCCAAAGTGGCGAGTTGACTTTACGAT 399  
DB 382 TACTTGATGATGTTGAAAAATTTATCATGCTCTCTGGAACCTGTGATTTGTAAT 441  
QY 400 CATTAGAGACACTAGAGCTAATATTAAAGTGGTTGMAAATGAAACACATGCCAA 459  
DB 442 GCAGCGATGAAGTTCGAAAGAGATTANTCTCGSTTGAAGCAAAACAAATGAAA 501  
QY 460 ATCAGAGAGTGTGTGAGGTGGCATAGCTCATCTGCTGTAATGTTGCTGTGAT 519  
DB 502 ATCAAGGACTTGTCCAGATGGCTCTATTAGTAGTCTTACCAAGCTGTGTGTGAAC 561  
QY 520 GCTGTGACTCTCAAGCAAGTGGCAATCAGCTTCCACAGAGCAAAACCAATATGC 579  
DB 562 ATGGTTATTTTAAAGGCATGAGCAGAGGATTTTAAAGAAAATATCAGAGAGAG 621  
QY 580 CATTCAATCTCCCAAGTCTCTGGAGGCAAGTCCCATGATGATCAGGAGCAAG 639  
DB 622 AAATTTTGGATGAATAGAGACAAAGTAAATCTGTACATGATGACACAGAGCAATCC 681  
QY 640 TTCAATTTGCTGTATTGAGGACCCATCATGAGATTTCTGAGCTCAGAT---ACAAT 696  
DB 682 TTATGCTTCACTTTCTGGAGACTTGAGGCCAAATTTCTAGGGATTCATATAAAAC 741  
QY 697 GTGGCATAAATCATGTACGTTCTGCTGCTCA-----GAATGACCTCTCTGAATTGA 750  
DB 742 AACGACCTTAACATGTTTGTGTTCTGCCAAGCATGATGCTGGAGAGATATA 801  
QY 751 ACAAACTGACCTTTGAGATCTAATGGAATGGAACCAATCCAAAGCAATGACCTTAAG 810  
DB 802 GATAAATAAGTCTGAGAAATTTGGTAGTGGACTAGTCCAGGGCATATGGAAGAAGA 861  
QY 811 TATGTGAGGTATTTTCTCAGTATGAGTGAAGAAATTTATGAATGAACATAT 870  
DB 862 AAGTGATCTGCTTGGCCGGTTTGGGTGAGGACAGTACGATCTTAGGCGGCTC 921  
QY 871 TTGAGAGCCCTAGGCTGAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGATT 930  
DB 922 CTGGTGCCTAGGGGATGGGATGCTCTTCTAGTGACCAAAAGCCGCTACTCTGGGATG 981  
QY 931 GCTTCGGGGGTGCTCTGTATATATCAAGGATGATGCAAAATCTTACATAGAGTCACT 990

DB 982 TCCTCAGGCTCCGGTGTGACGCCAGAAATTCCTGCACAGTTCCTTGTGGCAGTAACT 1041  
QY 991 GAGGAGGACCGAGGCTTACTGTGCGCCACAGGAAGTAAATTTAGAAAAGCAACTCCCT 1050  
DB 1042 GAGGAGGACCGAGGCTTACTGTGCGCCACAGTGCCTTACTGTACATCCGCCCA 1101  
QY 1051 CAGTCCACAGCTTTTAGAGCTGACCAACCTTCTTCTTATCAGGAAGATG----- 1105  
DB 1102 GGTCAAGAAATGTTCACTGCAATCATCCCTTCTCTTCTCATCAGGCAATGAATCC 1161  
QY 1106 -ACATCATCTTATCAGTGGCAAAGTTTCTTGCCTTGA 1143  
DB 1162 AACAGCATCTCTTCTTGGCAGATTTCTTCTCTCTTAA 1200

RESULT 13  
US-08-568-147B-1  
; Sequence 1, Application US/08568147B  
; Patent No. 5783422  
; GENERAL INFORMATION:  
; APPLICANT: Suminami, Yoshinori  
; APPLICANT: Kato, Hiroshi  
; APPLICANT: Sekiguchi, Kiyoshi  
; APPLICANT: Takeda, Katsumichi  
; TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL  
; TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/568,147B  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 800,952  
; FILING DATE: 02-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DiGiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8425  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1711 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 62..1234  
; US-08-568-147B-1

Query Match 18.5%; Score 211.6; DB 1; Length 1711;  
Best Local Similarity 54.5%; Pred. No. 1.8e-58;  
Matches 571; Conservative 0; Mismatches 444; Indels 33; Gaps 6;  
QY 1 ATGGCTCCCTGCTGCGAGCAATGCAGAGTTTGGCTTCAACCTGTCAGAGAGTGGAT 60

Db 62 ATGAATTCATCTCAGTGAAGCAACACCAAGTTCATGTTTCGACCTGTTCCACAGTTCCAGA 121  
Qy 61 GACATCAGGAAATGGAATGTTGTTCTTCTCTCTGAGCTCTTTCGCTCCCTCGCC 120  
Db 122 AAATCAAAAGAGAA---CAACATCTTCTATCTCCCTATCAGCATCATCAGCATAGGG 178  
Qy 121 CTGCTCCGCTTGGCGCTCAAGATGATCTCTCTCTCAGATTGATAGTTGTTCTCATGTT 180  
Db 179 ATGCTCTCTTAGAGCCAAAGACACACTGACACACAGATTAAAGAGGTTCTTCATTT 238  
Qy 181 -----AACCTGCTCAGGATGAGAACTCTCTTAATAGTCACTCAGG- 224  
Db 239 GATCAAGTCACAGAGAACACCAAGAAAGCTGCAACATATCATGTTGATAGTCCAGA 298  
Qy 225 --GCTCCAGTCTCACTGAAAGAGTTTTTCTGATTAATGCATCCACAGAGATAT 282  
Db 299 AATGTTTCATCACCAGTTTCAAAAGCTTCTGACTGAAATTCACAAATCCAGTATGCAT 358  
Qy 283 GATCTCAGCAATTCGAATGGCTTTTGTCTGAAAAGTGTATGGCTTTCATAAGGACTAC 342  
Db 359 GAGCTGAAGATGCCAACAGCTCTCGGAGAAAAGAGTATCTATTTTACAGGAATAT 418  
Qy 343 ATTGAGTGTCCGAAATATATACGATGCCAAAGTGGAGCGAGTTGACTTTACAGATCAT 402  
Db 419 TTAGATGCCATCAAGAAATTTTACAGACCGAGTGGAACTGTGATTTTGCATAATGCT 478  
Qy 403 TTAGAAGCACTAGACGTAATTAATTAATAGTGGTTGAAATGAAACACATGCGAAATC 462  
Db 479 CCAGAGAAAGTCCAAAGAGATTAACTCTGGTGGGAAAGTCAACGATGAAGAAATTT 538  
Qy 463 AAGAACGTGATGTTGGAAGTGGCATAGCTCATCTGCTGAATGTTGTTGTTGATGCT 522  
Db 539 AAAAACCTAATCTGAGGTAATATTGGCAGCAATACCATTTGTTCTTGTGAAGCA 598  
Qy 523 GTGACTTCAAGCGCAAGTGGCAATCAGCCCTTCAACAGACCGCAACCAATTAATTGCCAT 582  
Db 599 ATCTATTCAAGGGCGAGTGGGAGAAAGAAATTTAATAAGAGAGTACTAAGAGGAAAA 658  
Qy 583 TTCAAATCTCCCAAGTCTCTGGGAGGCGATGCGCATGATCAGGAAGGAGTTC 642  
Db 659 TTTTGGCCAAACAGAAATACATACAGTCCATACAGATGATGAGGCAATACATCTTT 718  
Qy 643 AATTTGCTGTTTATAGGACCCCATCAATGAAGATTCTTGAGCTCAGATCAATAGGT--- 699  
Db 719 CATTTTCTGCTGAGGATGATCAGGCCAAGTCTCTGGAATATACCATAAAGGCAAA 778  
Qy 700 GGCATAAACATGATGTTCTGCTGCTGAGA-----ATGACCTCTCTGAATTCGAAAC 753  
Db 779 GATCTAAGCATGATTTGTTGTGTCGCAATSAATAGATGTTCTTCCAGAGCTTGAAGAG 838  
Qy 754 AAACCTGACCTTCAGAAATCAATGAATGGACCAATCCAAAGGCGAATGACCTCTAAGTAT 813  
Db 839 AAACCTGCTGAGAAATGATGGAATGGACAGTTTGCAGATATGAGAGACAGT 898  
Qy 814 GTTAGGATTTTCTCAGTCAAGATAGAGAGATTAATGAATGAAGAAATATTTG 873  
Db 899 GTCGATTTTACACTTCTCGTCAAAAGTGAAGAGAGCTATGACCTTCAGGACACGTTG 958  
Qy 874 AGAGCCCTTAGGGCTGAAGATATCTTTGATGAATCCAGAGAGATCTCTCTGGATGCT 933  
Db 959 AGAACATAGGGAATGTTGATATCTCTAATG---GGGATGACAGACCTCTCAGGATGACC 1015  
Qy 934 TCGGGGGTCTCTGTATATATCAAGGATGATGCAAAATCTTACATAGAGTCACTGAG 993  
Db 1016 GGGAGCCGGTCTCTGTGCTATCTGGAGTCTTACACAGGGCTTTTGGAGGTTACAGAG 1075  
Qy 994 GAGGACCCGAGGCTACTGCTCCACAG 1021  
Db 1076 GAGGAGCAGAGGCTGAGCTGCCACCG 1103

RESULT 14

US-09-266-910-1

; Sequence 1, Application US/09266910  
; Patent No. 634362  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Use of a recombinant protein as receptor of a  
; TITLE OF INVENTION: Hepatitis virus  
; NUMBER OF SEQUENCES: 8  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/266,910  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1245 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Squamous Cell Carcinoma Antigen  
; TISSUE TYPE: Hepatoma  
; CELL TYPE: Hepatocyte  
; CELL LINE: HepG2  
; US-09-266-910-1  
  
Query Match 18.4%; Score 210; DB 4; Length 1245;  
Best Local Similarity 54.4%; Pred. No. 4.8e-58;  
Matches 570; Conservative 0; Mismatches 445; Indels 33; Gaps 6;  
  
Qy 1 ATGGCTCTCCCTCTGACACAAATGCAGAGTTTGTCTTCACTCTGTCAGAGAGATGAT 60  
Db 8 ATGAATTCATCTCAGTGAAGCAACCAAGTTCATGTTTCGACCTGTTCCACAGTTCCAGA 67  
Qy 61 GACATCAAGGAATGGAATGTTGTTCTCTCTCTGAGCTCTTTCGCTCCCTCGCC 120  
Db 68 AATCAAAAGAGAA---CAACATCTTCTATCTCCCTATCAGCATCATCAGCAATTAGGG 124  
Qy 121 CTGCTCCGCTTGGCGCTCAAGATGATCTCTCTCTCAGATTGATAGTTGTTCTCATGTT 180  
Db 125 ATGCTCTCTTAGAGCCAAAGACACACTGCAACAGATTAAGAGGTTCTTCACATTT 184  
Qy 181 -----AACCTGCTCAGGATGAGAACTCTCTTAATAGTCACTCAGG- 224  
Db 185 GATCAAGTCACAGAGAACACCAAGAAAGCTGCAACATATCATGTTGATAGGTCAGA 244  
Qy 225 --GCTCCAGTCTCAACTGAAAAGAGTTTTCGATATAAATGATCCCAAGAGTAT 282  
Db 245 AATGTTTCATCACCAGTTTCAAAGCTTCTGACTGAATTCACAAATCCACTGATCATAT 304  
Qy 283 GATCTCAGCATGTAAGTGGCTTTTGTCTGAAAAGTGTATGCTTTTCAAGAGACTAC 342  
Db 305 GAGCTGAAGATCGCCAAACAGCTCTTCGGAGAAAAAGCTATCTATTTTACAGGAATAT 364  
Qy 343 ATTGAGTGTCCGAAATTAATGATGCCAAAGTGGAGGAGTTGACTTTAGCATCAT 402  
Db 365 TTAGATGCCATCAAGAAATTTTACAGACAGTGTGGAAATCTGTTGATTTTGCATGCT 424  
Qy 403 TTAGAAGCACTCAGAGCTTAATTAATAGTGGGTTGAAAATGAAAACACATGCGAAATC 462  
Db 425 CCAGAGAAATTCGAAAGAGATTAACTCTCTGGTGGAAAGTCAACGATGAAGAAATTT 484  
Qy 463 AAGAACGTGATGTTGAGAGTGGCATAGCTCATCTGCTGTAATGTTGCTGGTGAATGCT 522  
Db 485 AAAACCTCTCTCTGAGGTAATTAATGGCAGCAATACCACATTTGTTCTTGTGAACGCA 544  
Qy 523 GTGACTTCAAGGCAAGTGGCAATCAGCTTCACAGAGGAGAAACCAATTAATTCAT 582







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2003, 23:40:42 : Search time 426 Seconds  
(without alignments)  
8917.544 Million cell updates/sec

Title: US-09-508-997A-1  
Perfect score: 1143  
Sequence: 1 atggctcccttgctgcagc.....gcaagtttttgccttga 1143

Scoring table: IDENTITY.NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1561799599 residues  
Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
13:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
14:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
15:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
16:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
17:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
18:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1143	100.0	1950	4	US-09-140-719-30
2	1143	100.0	1950	14	US-10-091-442-30
3	399.4	34.9	592	13	US-10-029-386-7445
4	399	34.9	399	13	US-10-029-386-7445
5	237.8	20.8	1200	11	US-09-823-187-1
6	224.8	19.7	1275	13	US-10-113-113-3
7	220.6	19.3	1279	15	US-10-157-031-354
8	220.6	19.3	1308	9	US-09-735-051-109
9	220.6	19.3	1308	10	US-08-850-716A-109
10	220.6	19.3	1308	10	US-09-897-778-109
11	220.6	19.3	1308	11	US-09-466-396A-109
12	220.6	19.3	1308	13	US-10-117-982-109
13	214.6	18.8	1539	9	US-09-764-898-25
14	211.6	18.5	1711	13	US-10-101-510-669
15	208.8	18.3	1193	8	US-08-731-566-1

16	207.8	18.2	1284	10	US-09-954-456-1841
17	206.8	18.1	1835	15	US-10-106-698-261
18	205	17.9	1419	9	US-09-735-705-111
19	205	17.9	1419	10	US-09-850-716A-111
20	205	17.9	1419	10	US-09-897-778-111
21	205	17.9	1419	11	US-09-466-396A-111
22	205	17.9	1419	13	US-10-117-982-111
23	204.6	17.9	2114	13	US-09-971-392-44
24	204	17.8	1897	15	US-10-208-408-17
25	204	17.8	1908	15	US-10-084-817-27
26	195.4	17.1	1883	10	US-09-895-814-944
27	195.4	17.1	1883	13	US-10-144-678A-944
28	195.4	17.1	1883	13	US-10-294-025-944
29	195.4	17.1	1883	14	US-10-012-886-944
30	194.4	17.0	1953	8	US-08-731-566-3
31	193.8	17.0	1549	15	US-10-073-979-12
32	188.4	16.5	1278	13	US-10-094-944-1
33	188.4	16.5	1624	10	US-09-925-300-500
34	187.2	16.4	1465	10	US-09-880-107-3027
35	187.2	16.4	1707	15	US-10-084-817-148
36	186.8	16.3	3603	13	US-10-113-113-1
37	184.4	16.1	1476	10	US-09-974-298-94
38	182.4	16.0	1230	15	US-10-267-311-54
39	182.4	16.0	2847	15	US-10-267-311-20
40	170	14.9	1626	13	US-09-993-363-1
41	158	13.8	1607	9	US-09-729-674-133
42	157.6	13.8	1819	13	US-09-993-363-3
43	156.8	13.7	2566	10	US-09-954-456-1813
44	156.8	13.7	2566	10	US-09-880-107-3283
45	155.2	13.6	1316	10	US-09-969-347-192

ALIGNMENTS

RESULT 1  
US-09-140-719-30  
; Sequence 30, Application US/09140719  
; Patent No. US20010026931A1  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Fuyuki  
; APPLICANT: TSUKUOKA, No. US20010026931A1UO  
; APPLICANT: NAKAZATO, Hiroshi  
; APPLICANT: MIURA, Kenju  
; APPLICANT: ISHIDA, No. US20010026931A1uhiro  
; APPLICANT: KURIHARA, Tatsuya  
; APPLICANT: YAMAUCHI, Kozo  
; APPLICANT: YAMAGUCHI, No. US20010026931A1cmi  
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/140,719  
; FILING DATE: 08-AUG-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER: US 08/474,661  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/091,028  
; FILING DATE: 14-JUL-1993

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1950 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; STRAIN: A431
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "DNA coding for human
; OTHER INFORMATION: megakaryocyte differentiation factor."
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 74..1217
US-09-140-719-30

Query Match 100.0%; Score 1143; DB 9; Length 1950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGCTGTCAGCAAAATCGAGAGTTTGGTTCAACCTGTTTCAGAGAGATGAT 60
DB 74 ATGGCTCCCTTGCTGTCAGCAAAATCGAGAGTTTGGTTCAACCTGTTTCAGAGAGATGAT 133
QY 61 GACAATCAAGAAATGGAATGTGTTTTCCTCTCTGAGGCTCTTCGGTGGCTGGCC 120
DB 134 GACAATCAAGAAATGGAATGTGTTTTCCTCTCTGAGGCTCTTCGGTGGCTGGCC 193
QY 121 CTGCTCCCTTGCGGCTCAAGATGACATCCCTCTCTCAGATGATGATGTTGTTTATGTT 180
DB 194 CTGCTCCCTTGCGGCTCAAGATGACATCCCTCTCTCAGATGATGATGTTGTTTATGTT 253
QY 181 AACACTGCTCAGGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 240
DB 254 AACACTGCTCAGGATATGGAATCTTCTTAATAGTCAGTCAGGCTCCAGTCTCAACTG 313
QY 241 AAAGAGTTTTTCTGATATAAATGATPCCCAAGGATATGATCTCAGCATTTGAT 300
DB 314 AAAGAGTTTTTCTGATATAAATGATPCCCAAGGATATGATCTCAGCATTTGAT 373
QY 301 GGGCTTTTGTGAAAAGGTATGGCTTTTCTAAGGCTACATGAGTGTGCGGAAA 360
DB 374 GGGCTTTTGTGAAAAGGTATGGCTTTTCTAAGGCTACATGAGTGTGCGGAAA 433
QY 361 TTATACGATGTCAAAGTGGAGGAGTGTGACTTTACGAATCATTTAGAGACACTAGAGT 420
DB 434 TTATACGATGTCAAAGTGGAGGAGTGTGACTTTACGAATCATTTAGAGACACTAGAGT 493
QY 421 AATATTAATAGTGGTGTGAAAATGAACACATGGCAAAATCAAGAACGTTGTTGAA 480
DB 494 AATATTAATAGTGGTGTGAAAATGAACACATGGCAAAATCAAGAACGTTGTTGAA 553
QY 481 GTGGCATAGCTCATCTGCTTAATGCTGTGCTGATGCTGTTGTTGTTTCAAGCAAG 540
DB 554 GGTGGCATAGCTCATCTGCTTAATGCTGTGCTGATGCTGTTGTTGTTTCAAGCAAG 613

```

```

QY 541 TGGCAATCAGCCTTCACCAAGGACGGAACCAATTAATTCGCAATTTCAATCTCCAGTGC 600
DB 614 TGGCAATCAGCCTTCACCAAGGACGGAACCAATTAATTCGCAATTTCAATCTCCAGTGC 673
QY 601 TCTGGGAAGGAGTTCGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 674 TCTGGGAAGGAGTTCGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
QY 661 GACCCATCAATGAAGATTTCTGAGCTCAGATACAAATGATGATGATGATGATGATGATGATGAT 720
DB 734 GACCCATCAATGAAGATTTCTGAGCTCAGATACAAATGATGATGATGATGATGATGATGATGAT 793
QY 721 CTGCTGAGATGACCTCTCTGAAATGAAACAAATGATGATGATGATGATGATGATGATGATGAT 780
DB 794 CTGCTGAGATGACCTCTCTGAAATGAAACAAATGATGATGATGATGATGATGATGATGATGAT 853
QY 781 TGGACCAATCCAGGGGAGATGACCTCTTAAGTATGTTGAGGATTTTTCCTCAGTTCAAG 840
DB 854 TGGACCAATCCAGGGGAGATGACCTCTTAAGTATGTTGAGGATTTTTCCTCAGTTCAAG 913
QY 841 ATAGAGAAAGATTTATGAATGAAACAAATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 914 ATAGAGAAAGATTTATGAATGAAACAAATGATGATGATGATGATGATGATGATGATGATGAT 973
QY 901 GATGAATCCAAAGCAGATCTCTCGGATGTTGCTTCGGGGGTGCTGCTGATATATCAAGG 960
DB 974 GATGAATCCAAAGCAGATCTCTCGGATGTTGCTTCGGGGGTGCTGCTGATATATCAAGG 1033
QY 961 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGGGCAACGAGGCTACTGCTGCACA 1020
DB 1034 ATGATGCAAAATCTTACATAGAGTCACTGAGAGGGGCAACGAGGCTACTGCTGCACA 1093
QY 1021 GGAAGTATATTTAGAAAGCAATCCCTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 1080
DB 1094 GGAAGTATATTTAGAAAGCAATCCCTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 1153
QY 1081 TTCTATTTTGTATCAGAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT 1140
DB 1154 TTCTATTTTGTATCAGAGGATGACATCATCTTATTCAGTGGCAAGTTTCTTGCCCT 1213
QY 1141 TGA 1143
DB 1214 TGA 1216

RESULT 2
US-10-091-442-30
; Sequence 30, Application US/10091442
; Publication No. US20020164711A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; INASA, Fuyuki
; TSURUOKA, No. US20020164711A1uo
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, No. US20020164711A1uhiro
; KURIHARA, Tatsuya
; YAMAGUCHI, Koza
; YAMAGUCHI, No. US20020164711A1omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/091,442  
FILING DATE: 07-Mar-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/140,719  
FILING DATE: 08-AUG-1998  
APPLICATION NUMBER: US 08/474,661  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
APPLICATION NUMBER: JP 6-067339  
FILING DATE: 04-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-247  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1950 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: A431  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1  
OTHER INFORMATION: /note= "DNA coding for human megakaryocyte differentiation factor."

FEATURE:  
NAME/KEY: CDS  
LOCATION: 74..1217  
SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
US-10-091-442-30

Query Match 100.0%; Score 1143; DB 14; Length 1950;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTCCCTTGTGCGGAAATGAGAGATTTGCTTCAACCTCTTCAGAGATGGAT 60  
Db 74 ATGGCTCCCTTGTGCGGAAATGAGAGATTTGCTTCAACCTCTTCAGAGATGGAT 133  
QY 61 GACATCAAGGAATGGAATGTTCTTTCCTCTCAGATTGATAGCTTCATGTT 180  
Db 134 GACATCAAGGAATGGAATGTTCTTTCCTCTCAGATTGATAGCTTCATGTT 193  
QY 121 CTGCTCCGCTGGGGCTCAAGATCACTCCCTCTCAGATTGATAGCTTCATGTT 180  
Db 194 CTGCTCCGCTGGGGCTCAAGATCACTCCCTCTCAGATTGATAGCTTCATGTT 253  
QY 181 AACATGCTCAGATATGGAATCTCTTAATAGTCACTCAGGCTCCAGTCTCACTG 240  
Db 254 AACATGCTCAGATATGGAATCTCTTAATAGTCACTCAGGCTCCAGTCTCACTG 313  
QY 241 AAAAGAGTTTTTCTGATATAAATGATCCACAGGATATGATCAGATTGGAT 300  
Db 314 AAAAGAGTTTTTCTGATATAAATGATCCACAGGATATGATCAGATTGGAT 373  
QY 301 GGGCTTTTGTGAAAAAGTATGGCTTTTCAAGGACTACATTGAGTGGCGGAAA 360  
Db 374 GGGCTTTTGTGAAAAAGTATGGCTTTTCAAGGACTACATTGAGTGGCGGAAA 433  
QY 361 TTATCAGATGCCAAAGTGGAGGAGTTGACTTTACGAATCAATTAGAGACACTAGAGT 420

Db 434 TTATACGATGCCAAAGTGGAGGAGTTGACTTTTACGAATCAATTTAGAGACACTAGAGT 493  
QY 421 AATATTAATAGTGGGTTCAAAATGAAACACATGGCAAAATCAAGACGATGTTGGTCAA 480  
Db 494 AATATTAATAGTGGGTTCAAAATGAAACACATGGCAAAATCAAGACGATGTTGGTCAA 553  
QY 481 GTGGCATAGAGTCATCTCTGTATATGGTCTGGTGAATGCTGTACTTCAAGAGCAAG 540  
Db 554 GTGGCATAGAGTCATCTCTGTATATGGTCTGGTGAATGCTGTACTTCAAGAGCAAG 613  
QY 541 TGGCAATCAGCTTCCAGAGAGGAAACATTAATTTGCCATTTCCAAATCTCCAGAGTC 600  
Db 614 TGGCAATCAGCTTCCAGAGAGGAAACATTAATTTGCCATTTCCAAATCTCCAGAGTC 673  
QY 601 TCTGGGAAGCAGTCGCCATGATGATCAGGAACGGAGATTTCAATTTGCTGTATTTCAG 660  
Db 674 TCTGGGAAGCAGTCGCCATGATGATCAGGAACGGAGATTTCAATTTGCTGTATTTCAG 733  
QY 661 GACCATCAATGAAGATTTCTTGAAGTCTCAGATCAATATGGTGGCATAAATGATGTCG 720  
Db 734 GACCATCAATGAAGATTTCTTGAAGTCTCAGATCAATATGGTGGCATAAATGATGTCG 793  
QY 721 CTGCTGAGAAATGACCTCTCTGAAATTTGAAACAACTGACCTTCAGAAATCTAATGAA 780  
Db 794 CTGCTGAGAAATGACCTCTCTGAAATTTGAAACAACTGACCTTCAGAAATCTAATGAA 853  
QY 781 TGGACCAATCCAAAGCGAATGACCTCTTAAGTATGTTGAGGTATTTTCTCAGTTCAG 840  
Db 854 TGGACCAATCCAAAGCGAATGACCTCTTAAGTATGTTGAGGTATTTTCTCAGTTCAG 913  
QY 841 ATAGAGAAGAAATTAAGAAATGAAACAAATTTTACAGAGCCCTTAGGGCTGAAGATATCTTT 900  
Db 914 ATAGAGAAGAAATTAAGAAATGAAACAAATTTTACAGAGCCCTTAGGGCTGAAGATATCTTT 973  
QY 901 GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTCGGGGGGTGTCTGTATATATCAAG 960  
Db 974 GATGAATCCAAAGCAGATCTCTCTGGGATTTGCTTCGGGGGGTGTCTGTATATATCAAG 1033  
QY 961 ATGATGCACAAAATCTTACATAGAGTCACTGAGGAGGGCACCGAGGCTACTGTCGCCACA 1020  
Db 1034 ATGATGCACAAAATCTTACATAGAGTCACTGAGGAGGGCACCGAGGCTACTGTCGCCACA 1093  
QY 1021 GGAAGTATATATGTAAGAAAGCAATCTCCCTCAGTCCAGCTGTTTAGAGCTGACACCCA 1080  
Db 1094 GGAAGTATATATGTAAGAAAGCAATCTCCCTCAGTCCAGCTGTTTAGAGCTGACACCCA 1153  
QY 1081 TTCTATTGTTTATCAGGAAGGATGACATCATCTTTATTCAGTGGCAAGTTTCTTGCCT 1140  
Db 1154 TTCTATTGTTTATCAGGAAGGATGACATCATCTTTATTCAGTGGCAAGTTTCTTGCCT 1213  
QY 1141 TGA 1143  
Db 1214 TGA 1216

## RESULT 3

US-10-029-386-7445  
; Sequence 7445, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AFOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 7445  
; LENGTH: 592  
; TYPE: DNA

```

RESULT 4
US-10-029-386-21145
/ Sequence 21145, Application US/10029386
/ Publication No. US20030194704A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Shaaron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
/ TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
/ FILE REFERENCE: AEOmica-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 21145
/ LENGTH: 399
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR18.3
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6

```

RESULT 5

US-09-823-187-1

Sequence 1, Application US/09823187

Publication No. US20030096952A1

GENERAL INFORMATION:

APPLICANT: Burgess, Catherine

APPLICANT: Gusev, Vladimir Y

APPLICANT: Liu, Xiaohong

APPLICANT: Majumder, Kumud

APPLICANT: Padigar, Muralidhar

APPLICANT: Patturajan, Meera

APPLICANT: Shimketa, Richard A

APPLICANT: Spaderna, Steven K

APPLICANT: Spytek, Kimberly

APPLICANT: Taupier, Raymond J

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: 15966-745

CURRENT APPLICATION NUMBER: US/09/823,187

CURRENT FILING DATE: 2001-03-29

PRIOR APPLICATION NUMBER: 60/193,339

PRIOR FILING DATE: 2000-03-30

PRIOR APPLICATION NUMBER: 60/193,205

PRIOR FILING DATE: 2000-03-30

PRIOR APPLICATION NUMBER: 60/195,343

PRIOR FILING DATE: 2000-04-05

PRIOR APPLICATION NUMBER: 60/195,088

PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: 60/195,005

PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: 60/195,792

PRIOR FILING DATE: 2000-04-10

PRIOR APPLICATION NUMBER: 60/196,556

PRIOR FILING DATE: 2000-04-11

PRIOR APPLICATION NUMBER: 60/197,081

;; PRIOR FILING DATE: 2000-04-13  
;; PRIOR APPLICATION NUMBER: 60/197,525  
;; PRIOR FILING DATE: 2000-04-14  
;; PRIOR APPLICATION NUMBER: 60/197,087  
;; PRIOR FILING DATE: 2000-04-14  
;; NUMBER OF SEQ ID NOS: 103  
;; SOFTWARE: Patent In Ver. 2.1  
;; SEQ ID NO 1  
;; LENGTH: 1200  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (7)..(1191)  
US-09-823-187-1

Query Match 20.8%; Score 237.8; DB 11; Length 1200;  
Best Local Similarity 52.9%; Pred. No. 6.3e-60;  
Matches 608; Conservative 0; Mismatches 502; Indels 39; Gaps 3;  
  
QY 1 ATGGGCTCCCTTCTGCGCAAAATGCGAGTTTTCCTTCAACTGTTTCAGAGAGATGGAT 60  
DB 7 ATGGACTCTCTCTTACAGCAACACCAAAATTTTCCTTTCATCTTTTCAGAGATAGG 66  
  
QY 61 GACAAATCAGGAATGGAATGTTCTTTCTCTCTGAGGCTCTTCTGCTGCCCTGCC 120  
DB 67 AAAGATGATCGTCATAAAACAVATTTTCTCTCCCTGAGGCTCTCAGCTGCCCTGGT 126  
  
QY 121 CTGCTCGGCTTGGGCTCAGATGACTCCCTCTCAGATTGATAGTGTCTCATGTT 180  
DB 127 ATGTAGCTTGGTCTAGAGTGAGTGACATGACATGATGATGAGTACTACTACTC 186  
  
QY 181 AACACTGCTCTCAGGATATGGAATCTTCTTAATAGTCAGTC-----A 222  
DB 187 AACGAATTTCCCAAGTAAAGCAAGAACCTCTGGTCTTAAACAAATGAGAGCGGA 246  
  
QY 223 GGGCTCAGTCTCACTGAAAGAGTTTTCCTGATATAAATCAATCCCAAGGATTAT 282  
DB 247 CTGCTCAGCTGCTACTTTGGGCACTTCTCTCCAAATTAGACAGGATCAAGACTGAT 306  
  
QY 283 GATCTCAGCATTGTAATGGGCTTTTCTGTAAGAGTGTATGGCTTTCATAGGACTAC 342  
DB 307 ACCTGAGTATTGCCACAGGCTTTATGAGAGAGGAATTCCTCAATCTGTGAGGAATAC 366  
  
QY 343 ATTGAGTGTCCGAAATATATACGATGCCAAAGTGGAGCGAGTGTACTTTACGAATCAT 402  
DB 367 TTAGATGTTGATTCATTTTACCAACAGCAATGAAAGTGTGATTTCCAAATAAC 426  
  
QY 403 TTAGAAGACACTAGACGTAATATTAATAGTGGGTGAAATGAACACATGSCAAATC 462  
DB 427 CCTGAAATAATCCAGCAAGAGATTAATCTTGGGTGAAATGTCATCCCAAGGTAAATC 486  
  
QY 463 AAGAAGCTGATTTGGTGGAGTGGCATAGCTCTCTCTGATGATGGTGTGATGATGT 522  
DB 487 AAGGACTCTTTCAGACAGGAGCTTATATGCTGAGACGTGCTGTACTGTGATGATGT 546  
  
QY 523 GTGTATTTCAAAGGCAAGTGGCAATCAGCCTTCCCAAGAGCAAAACCATTAATTCGCAAT 582  
DB 547 GTTACTTCAAGGCCAATGGGAACATATCTTTCACCAAGAAACACGCTGGTGTACCT 606  
  
QY 583 TCCAAATCTCCCAAGTGTCTGGGAAGGCAATGCCCATGATGATCAGAACGGAAGTTC 642  
DB 607 TCTGTCTAAATTCAGATGAAACAGAGTGTGAAGATGATGACGCAAAAGGCTCTAC 666  
  
QY 643 AATTTCTGTTTATTTAGGACCCATCAATGAGATTTCTGAGTCAAGTA---CAATGTT 699  
DB 667 AGAATTTGCTTATAGAGGAGGTGAAGGCAAGATCTCTGGAATGAGGTACCAAGGGG 726  
  
QY 700 GGCATAAACATGATGTTCTGCTGCC-----TGAGATGACCTCTCT 741  
DB 727 AAGCTCAGCATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786  
  
QY 742 GAAATGAAACAACTGACCTTTTCAGAACTTAATGGAATGGACCAATCCAGGCGAATG 801

DB 787 GAGCTTGAAGGAAAAATCACCTATGAAAAATGTTGGCTGGAGCAGCTCAGAAAAATG 846  
QY 802 ACCTCTAAGTATGTTGAGGTATTTTTCCTCAGTCAAGATAGAGAAATTTATGAATG 861  
DB 847 TCAGAAGATCGGTGGTCTCTCTTCCCGGGTTCCCTTGGAGACAGCATGATCTC 906  
QY 862 AAACAATTTTGAAGCCCTAGGCTGAAAGATATCTTTGATGATCCAAAGCAGATCTC 921  
DB 907 AATTCCTATTTTACAAGACATGGGCAATACGATATCTTTGATGAAACGAGGCTGATCT 966  
QY 922 TCTGGGATTTCTCGGGGGTCTGCTGATATATATCAAGGATGATGCAAAATTTACATA 981  
DB 967 ACTGAAATCTCTCAAAGTCCCAATTTGTACTTTGTCAAAATTTATCCCAAACTTTGTG 1026  
QY 982 GAGTCACTGAGGAGGACCGAGGCTACTGCTGCCACAGGAAGTAAATTTAGAAAG 1041  
DB 1027 GAGTGGATGAAACGATACCCAGGACGCTGACGACCTGGGGCTGTTGCTCGAAAGG 1086  
QY 1042 CAACTCCCTCAGTCCAGCTGTTTAGAGCTGACCAACCATCTTATTTGATCAGGAAG 1101  
DB 1087 TCACCTACGATCTTGGGTGGAGTTTAATGCCAACCAACCTTTCTCTTTTCATTAGACAC 1146  
QY 1102 GATGACATC 1110  
DB 1147 AACAAACC 1155  
  
RESULT 6  
US-10-113-113-3  
; Sequence 3, Application US/10113113  
; Publication No. US20030166852A1  
; GENERAL INFORMATION:  
; APPLICANT: Holloway, James L.  
; TITLE OF INVENTION: PROTEINASE INHIBITOR ZSERP9  
; FILE REFERENCE: 00-92  
; CURRENT APPLICATION NUMBER: US/10/113,113  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: 60/280,678  
; PRIOR FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1275  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: degenerate sequence  
; NAME/KEY: misc feature  
; LOCATION: (1)-(1275)  
; OTHER INFORMATION: n = A, T, C or G  
US-10-113-113-3

Query Match 19.7%; Score 224.8; DB 13; Length 1275;  
Best Local Similarity 37.2%; Pred. No. 4.9e-56;  
Matches 328; Conservative 154; Mismatches 379; Indels 21; Gaps 2;  
  
QY 248 TTTTCTGATATAAATCCATCCCAAGGATTTATGATCTCAGCATTTGTAATGGCTTT 307  
DB 356 TNYTWSNAARYTNGAYMGNATHAAACNAYTACNYTWSNATHGNCNAYMGNATNT 415  
  
QY 308 TTGCTGAAAGAGTGTATGGCTTTCATAGGACTACATTTGAGTGTGCCGAAAAATTAACG 367  
DB 416 AYGNHARCARARTYTCNATHTGTCARGARTAYTNGAYGNGNATHTHCARTTYTAYC 475  
  
QY 368 ATGCCAAAGTGGAGGCTGACTTTTACGAATCATTTAGAGACACTAGAGTATATTA 427  
DB 476 AYACNACNATHGARNSTNGAYTTCARARAYCCNHAARPNMGNMGNRARGARATHA 535  
  
QY 428 ATAAGTGGGTGAAATGAACACATGCGCAAAATCAAGAACCTGATTTGTTGAGGCTGGCA 487  
DB 536 AYTTTGGGTNGARTGYCARWSNCARGGNAARATHAARGARYTNTTYSNARAGYCNRA 595









Db 502 ATCAAGGACTTGTCCCGATGGCTCTATTAGTAGCTCTACCAAGCTGGTGTGTAAC 561  
Qy 520 GCTGTGTACTTCAAAAGGCAAGTGGCAATCAGCTTTCACAGAGCCGAAACCTAATTCG 579  
Db 562 ATGGCTTTATTTTAAAGGCAATGGACAGGAGTTTAAAGAAAGAAATCTAAGGAAG 621  
Qy 580 CATTTTAAATTCCTCCAAAGTGTCTGGGAAGGAGTGGCCATGATGATCAGGAACGGAG 639  
Db 622 AAATTTTGGATGAATAAGAGCACAAAGTAATCTGTACAGATGATGACACAGCCATTC 681  
Qy 640 TTCAATTTGTCTGTATTGAGGACCCATCAATGAGAGATCTTGAGCTGAGT---ACAAT 696  
Db 682 TTATAGCTTCACTTTCTCGAGGACTTGCAGGCCAAAATCTAGGGATTCCATATAAAAC 741  
Qy 697 GGTGGCAAAAATCATAGCTTCTGTCTGCTGCTGA-----GAATGAOCTCTCTGAAATGAA 750  
Db 742 AACGACCTAAGCAATTTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801  
Qy 751 AACAACTGACCTTTTCAAGATCTTAATGGAATGGACCAATCCAAAGGCGAATGACCTTAAG 810  
Db 802 GATAAATAAAGTCTCGAATTTGGTAGAGTGGACTAGTCCAGGGCATATGAGAAAGA 861  
Qy 811 TAGTGTGAGGATTTTCTCTCAGTTTCAGATAGAGAAATATGAAATGAAACAAATAT 870  
Db 862 AAGTGGAATCTGCACTTGGCCCGGTTGAGGTGGAGGACAGTTACGATCTAGAGCGGTC 921  
Qy 871 TTGAGAGCCCTAGGCTGAAAGATATCTTTGATGAATCCAAAGGCGAATGACCTTAAG 930  
Db 922 CTGGCTGCCATGGGATGGCGATGCTTCAAGTGGCTCTTAATGAGTGGTCTGCTGCTGCT 981  
Qy 931 GCTTGGGGGTGCTCTGTATATATCAAGGATGATGCAAAATCTTACATAGAGGTCAT 990  
Db 982 TCCTCAGGCTCCGGTTGTACGCGCAGAAAGTTCCTCAGAGTTCTCTTGTGGCAGTAAT 1041  
Qy 991 GAGGAGGCGACCGAGGCTACTGCTGCCACAGGAAGTAAATTTAGTAAAGCAATTC 1050  
Db 1042 GAGGAGGCGACCGAGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101  
Qy 1051 CAGTCCACGCTGTTTAGAGCTGACCAACCATCTTATTTGTTATCAGGAGGATG---- 1105  
Db 1102 GGTATGAAATGTTCACTGCAATCATCCCTTCTGTTCTTCAATGAGCAGATGAATCC 1161  
Qy 1106 -ACATCACTTATTCAGTGGCAAGTTCCTTGGCCCTTGA 1143  
Db 1162 AACAGCATCTCTTCTCGGCAAGTTTCTTCTCTCTTAA 1200

RESULT 11  
US-09-466-396A-103  
; Sequence 109, Application US/09466396A  
; Publication No. US20030119763A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C4  
; CURRENT APPLICATION NUMBER: US/09/466.396A  
; CURRENT FILING DATE: 1999-12-17  
; NUMBER OF SEQ ID NOS: 224  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 109  
; LENGTH: 1308  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-466-396A-109

Query Match 19.3%; Score 220.6; DB 11; Length 1308;  
Best Local Similarity 52.7%; Pred. No. 9e-55;  
Matches 621; Conservative 0; Mismatches 519; Indels 39; Gaps 5;

Qy 1 ATGGCTCCCTTGTGTCAGCAATGACAGATTTTGTCAACCTGTTTACAGAGATGAT 60  
Db 25 ATGGATTCATTTGGCCCGGCTGACACTCGACTGGGTGTGATCTTTTCAAGAGCTGAAG 84



```
QY 310 GCTGAAAGTGTATGCTTTCATAGAGCTACATTGAGTGTGCGAAATATACGAT 369
Db 177 GGGACAAAGACGATGGCATTTTCATAGCAATATTTAAGCTGTCTGAGAAATGGTATCAA 236
QY 370 GCCAAGTGGAGCGAGTGTACCTTTACGAATCATTTAGAAAGACATAGACGTATATTAAT 429
Db 237 GCCAGTTTGCAATCTGTGGATTTTGAACAGCTCTACAGAGAAACGAGGAAACGATTAT 296
QY 430 AAGTGGTTGAAATGAAACATACATGCGAATATCAAGACGTCATGTTGAGTGGCATA 489
Db 297 GCTTGGTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 356
QY 490 AGCTCATCTGCTGTATGTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 549
Db 357 GACCTTTCATCTGTATGTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 416
QY 550 GCTTTCACCAAGAGCGAAACCAATATTTGCCATTTCCAAATCTCCCAAGTCTCTGGGAG 609
Db 417 AAATTTCAAGTAAAGAGACAGTATAAAGTCTTTTACGTAAGTGAAGTAAATGTA 476
QY 610 GCAGTGGCATGATGATCAGGAACGGAAGTCAATTTCTGTGTTATGAGGACCCATCA 669
Db 477 ACTGTGGAATGATGATCAGGAACGGAAGTCAATTTTAACTGGCTTTTGAAGGAGCGGAG 536
QY 670 ATGAAGATCTTGTAGCTCAGATACATGCTGGCATTAACATGCTAGCTTCTGCTGTGCTGAG 729
Db 537 ATGCAAGTCTTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 596
QY 730 AATGACCTCTCTG-----AAATTGAAACAAACCTGACCTTTCAGAAATCTTAATGGAA 780
Db 597 GTAGCATAGCTAATCTGAAACAGATAGAAAGCAGCTGAATTCGGGAGCTTTTCATGAG 656
QY 781 TGGACCAATCCAGGCGATGACCTTAAGTATGTTGAGTATGTTTTCCTCAGTTCAAG 840
Db 657 TGGACCAAGCTCTTCAACATGATGGAAGAGGAAGTGAAGTACACCTCCCGAGATTCAA 716
QY 841 ATAGAGAAAGATATGAAATGAAACATATTTGAGAGCCTGAGGCTGGAAGATATCTTT 900
Db 717 CTGAAATAGTATGCTAATTTCCCTGTTTAAACCTCCAGGGGTGACAGATCTCTTC 776
QY 901 GATGAATCCAAAGCAGATCTCTCTGGATTTGCTTGGGGGGTCTGTGTATATATCAAG 960
Db 777 AACCAAGTCAAAAGCTGATCTTCTGGAATGTCACCAACCAAGGCTTATATTTATCAAA 836
QY 961 ATGATGCAAAATTTACATAGAGTCTGCTGAGGAGGCGACGAGGCTACTGCTGCA 1020
Db 837 GCATCCCAAGTCTACTCTGATGTACGGAAGAGGCGACGAGGCGAGCAGCCACT 896
QY 1021 GGAAGTAAATTTAGAAAGCAACTCCCTCAGTCCACGCTGTTTGAAGTCAACCCCA 1080
Db 897 GGGACAGCATCTGTTTAAAGCTTACCAATGAGAGCTCAGTTCAAGGGCAACCCCCC 956
QY 1081 TTCCATTGTTTATCAGGAAGGATGATCATC 1113
Db 957 TTCCGTTCTTTTAAAGGACACTCATACCAAC 989
```

## RESULT 14

US-10-101-510-669

; Sequence 669, Application US/10101510

; Publication No. US20030148295A1

; GENERAL INFORMATION:

; APPLICANT: WAN, JACKSON

; APPLICANT: WANG, YIXIN

; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE

; FILE REFERENCE: 15117.0012

; CURRENT APPLICATION NUMBER: US/10/101,510

; PRIOR FILING DATE: 2002-03-20

; PRIOR APPLICATION NUMBER: 60/276,947

; PRIOR FILING DATE: 2001-03-20

; NUMBER OF SEQ ID NOS: 805

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 669

```
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-669
```

## Query Match

18.5%; Score 211.6; DB 13; Length 1711;

Best Local Similarity 54.5%; Pred. No. 5.2e-52;

Matches 571; Conservative 0; Mismatches 444; Indels 33; Gaps 6;

```
QY 1 ATGGCTCCCTGCTGTGAGCAAAATGACAGATTTTGTTCAACTGTTCCAGAGAGATGGAT 60
Db 62 ATGANTTCACTCACTGAGCAAGCCACACCAAGTTTCATGTTGACCTGTTCCACAGTTTCA 121
QY 61 GACATCAAGGAATGGAATGTTTTCCTCTCTCTGAGCCTCTTCGTCGCCCTGGCC 120
Db 122 AAATCAAAAGAA---CAACATCTTCTATTCCTCCCTATCAGCATCACATCAGCATTAGGG 178
QY 121 CTGGTCCGCTTGGGCGCTCAGATGATGATCTCTCTCAGATTCATGATGATGATGATGAT 180
Db 179 ATGGTCTCTTAGGAGCCAAAGACCAACTGCAACACAGATTAAAGAGGTTCTTCACTTT 238
QY 181 -----AACACTGCTCAGGATATGAAACTCTTCTTAAATAGTCAGTCAGG- 224
Db 239 GATCAAGTCACAGAGAACACCCAGGAAAGCTGCAACATATCATGTTGATAGTCAGGA 298
QY 225 --GCTCCAGTCTCAACTGAAAGAGTTTTTCTGATATAAATGCAATCCCAAGGATTAT 282
Db 299 AATGTTTCATCACCAGTTTCAAAGCTTCTGACTGAAATTCACAAATCCACTGATGATAT 358
QY 283 GATCTCAGCATTTGATGGCTTTTCTGTAAGAGTGTATGGCTTTTCATAGGACTAC 342
Db 359 GACTGAGATCGCCAAAGCTCTTCGGAGAAACAAACGTATCTATTTTACAGGAATAT 418
QY 343 ATTGACTGTGGCGAAAAATATACGATCCAAAGTGGAGCGAGTTGACTTTTACGAATCT 402
Db 419 TTAGTGGCATCAAGAAATTTACAGACACAGTGTGGAATCTGTTGATTTGCAATGCT 478
QY 403 TTAGAGACACTAGACCTATATTAATGAGTGGTTGAAATGAAACACATGCGAAAAATC 462
Db 479 CCAGAGAAAGTTCGAAAGAGATTAATCTCTGGGTGAAAGTCAACGAATGAAAAAT 538
QY 463 AAGACGTGATTTGTTGAGGTGGCATAGCTATCTCTGTTAATGGTCTGTGTTGTTGTTGCT 522
Db 539 AAACACCTTATTTCTGAGGTATATTTGGAGCAATACCACTGTTGTTGTTGTTGTTGTTG 598
QY 523 GTGATCTTCAAAGCAAGTGGCAATCAGCTTTCACAGAGCGAAACCATATAATTTGCCAT 582
Db 599 ATCTATTTCAAAGGCGAGTGGGAGAGAAATTTAATAAAGAGATCTATAAGAGGAAAAA 658
QY 583 TTCAATCTCCCAAGTCTCTGGGAGGCGAGTGGCAATGATGATGATGATGATGATGATGAT 642
Db 659 TTTTGGCCAAAACAAAGATACATCAAGTCCATACAGATGATGAGCAATACATCTTTT 718
QY 643 AATTTCTCTGTTTATGAGGACCCATCAATGAGATCTTGGCTCAGATCAATGTT--- 699
Db 719 CATTTTGGCTCGCTGGAGAGTGTACAGCCAGGTCTCTGMAATACCATCAAGGCAAA 778
QY 700 GGCATMAACATGTACGTTCTGCTGCTGAGA-----ATGACCTCTCTGAAATTTGAAAC 753
Db 779 GATCTAAGCATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 838
QY 754 AAATGACCTTTTCAAGATCTAATGGAATGAGCAATCCAGGCGAATGACCTCTAAGTAT 813
Db 839 AAATGACCTGCTGAGAAATGATGGAATGGAAGTTTTCAGAAATATGAGAGGACACGT 898
QY 814 GTTGAAGTATTTTCTCAGTTTCAAGATGAGAGAAATTTATGAAATGAAACATATTTTG 873
Db 899 GTCGATTACCTTACCTCGTTTCAAGTGGAGAGAGCTATGACCTTCAAGGACACGTTG 958
QY 874 AGAGCCCTAGGCGTGAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATGCT 933
Db 959 AGAACCCATGGGAATGGTGGATATCTTCAATG---GGGATGCGAGCTCTCAGGCAATGACC 1015
```

QY 934 TCGGGGGTCTCTGTATATCAAGGATGATGCAAAATCTTACATAGAGTCACTGAG 993  
Db 1016 GAGAGCCGGTCTCGTGTATCTGAGTCTACACAAGGCTTGTGGAGTTACAGAG 1075  
QY 994 GAGGACCCGAGGCTACTGCTGCCACAG 1021  
Db 1076 GAGGAGCAGAGCTGAGCTGCCACCG 1103

## RESULT 15

US-08-731-566-1  
; Sequence 1, Application US/08731566  
; Publication No. US20030072752A1  
; GENERAL INFORMATION:  
; APPLICANT: Toshihiko Hibino et al.  
; TITLE OF INVENTION: TWO NOVEL GENES FROM PSORIATIC EPIDERMIS:  
; TITLE OF INVENTION: PSORIASITIN TYPE I AND PSORIASITIN TYPE II  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1775  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/731,566  
; FILING DATE: 16-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/013,755  
; FILING DATE: 20-MAR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/005,679  
; FILING DATE: 17-OCT-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Louis Myers  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: MGP-040CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1193 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 22..1193  
US-08-731-566-1

Query Match 18.3%; Score 208.8; DB 8; Length 1193;  
Best Local Similarity 53.7%; Pred. No. 2.8e-51;  
Matches 563; Conservative 0; Mismatches 452; Indels 33; Gaps 5;

QY 1 ATGGCTCCCTTGGTCGAGCAATGCAAGATTTTGGTTCACCTGTTACAGAGATGGAT 60  
Db 22 ATGAATCTCAGTGAAGCCACACCAAGTTTCATGTCGACCTGTTCCACAGTTCA 81  
QY 61 GACAAATCAAGGAATGAAATGTTCTTTCTCTGAGGCTCTTGCTGCCCTGGGC 120  
Db 82 AAATCAAAAGAA--CAACATCTTCTATTCCTATCAGCATCAGATCAGATTAGG 138  
QY 121 CTGTCGGCTGGGGCTCAAGATGACTCCCTCTCTCAGATGCAATGATTCCTCATGTT 160  
Db 139 ATGGTCTCTTAGGAGCCAAAGGCAACACTGCAACAAGATTAAAGAGTTCTTCACTTT 198

QY 181 AACATGCTCAGGATATGGA-----ACTTCTTATAGTCACTCA 222  
Db 199 GATCAAGTCAAGAGAACACCAAGAAAAGCTGCAACATATCATGTGTATAGTCA 258  
QY 223 GGGCTCCAGTCTCAACTGAAAGAGTTTTTCTGATATAATGCAATCCCAAGGATTA 282  
Db 259 GATGTTATCACCAGTTTCAAGGCTTCTGACTGATTCACAAATCACTGATGATAT 318  
QY 283 GATCTCAGCATTTGAAATGGGCTTTTCTGAAAAGTGTATGGCTTTTATAAGGACTAC 342  
Db 319 GAGCTGAGATGCGCAACAAGCTCTCGGAGAAAACGATATCTATTTTACAGGAATAT 378  
QY 343 ATTGAGTGTCCGAAAATTTATACGATGCCAAAAGTGGAGCGGTGACTTACGAATCAT 402  
Db 379 TTAGATGCCATCAAGAAATTTTACAGACCAGTGTGGAATCTGTGATTTGCAATGCT 438  
QY 403 TTAGAAGACACTAGACGCTAATTAATAAGTGGGTGAAAATGAAACACATGGCAAAATC 462  
Db 439 CCAGAGAAGTCCGAAAAGATTAATCTCTGGGTGMAAGTCAACCGAATGAAAATTT 498  
QY 463 AAGACGTGATTGTGAAGTGGCATPAAGCTCATCTGCTGTAATGCTGCTGGTGAATGCT 522  
Db 499 AAAAACTAATTCCTGAAGGTAATATGGCAGCAATACCAATGCTTCTTCTGTAAGCA 558  
QY 523 GTCTACTTCAAGAGCAAGTGCATCAGCCTTCCAAAGAGCGMAACCATAAATGSCAT 582  
Db 559 ATCTATTTCAAGGGCAGTGGGAGAGAAATTTAATAAGAGATCTAAGAGGAAAAA 618  
QY 583 TTCAATCTCCCAAGTGTCTCTGGAAGGAGTGCCTCCCATGATGCATCAGGAACGGAAGTTC 642  
Db 619 TTTTGGCCAAACAAGAAATACATACAGTCCATACAGATGATGAGGCAATACACATCTTTT 678  
QY 643 AATTGTCTGTTATGAGGACCCATCATGAAGATTTCTTGAGCTCAGATCAATGCT--- 699  
Db 679 CATTTTGGCTCGCTGGAGGATGTACAGGCCAAGGTCTCTGMAATACCATACAAAGCAAA 738  
QY 700 GGCATAAATCATGTCGTTCTGCTGCTGAGA-----ATGACCTCTCTGAAATGAAAC 753  
Db 739 GATCTAAGCATGATTGTGTTGTTGTCGCCAATGAAATCGATGTTCTCAGAGGCTTGAAGAG 798  
QY 754 AACTGACCTTTTCAAGAAATCTAATGGAATGGACCAATCCAAAGCGAATGACCTCTAAGTAT 813  
Db 799 GAATCACTGCTGAGAAATTTGATGGAATGGACAAGTTTGCAGAAATATGAGAGACACGT 858  
QY 814 GTTGAAGTATTTTCTCTCAGTTCAAGATAGAGAGAAATTTATGAATGAAACATATTG 873  
Db 859 CTCGATTTTACCTTACTCTCGTTCAAAGTGGAGAGAGCTATGACCTCAAGGACACGTTG 918  
QY 874 AGAGCCTTAGGGCTGAAAGATATCTTTGATGAATCCAAAGCAGATCTCTCTGGGATGCT 933  
Db 919 AGAATCATGGAAATGTTGATATCTTCAATG---GGGATCGAGACCTCTCAGGCTAGACC 975  
QY 934 TCGGGGGGTCTCTGTATATATCAAGATGATGCAAAATCTTACATAGAGTCACTGAG 993  
Db 976 GGGAGCCGGTCTCTGCTCTATCTGAGTCTGCAAGGCTTTGTGGAGTTTACAGAG 1035  
QY 994 GAGGCACTGAGGCTACTGCTGCCACAG 1021  
Db 1036 GAGGAGCAGAAAGCTGAGCTGCCACCG 1063

Search completed: December 14, 2003, 01:13:19

Job time : 437 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2003, 00:59:44 ; Search time 2773 Seconds  
(without alignments)  
10018.037 Million cell updates/sec

Title: US-09-508-997A-1  
Perfect score: 1143  
Sequence: 1 atggcctcccttgctgcgc.....gcaagtttttgccttgga 1143

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1:	em_estba:**
2:	em_esthum:**
3:	em_estin:**
4:	em_estmu:**
5:	em_estov:**
6:	em_estpl:**
7:	em_estro:**
8:	em_htc:**
9:	gb_est1:**
10:	gb_est2:**
11:	gb_est3:**
12:	gb_est4:**
13:	gb_est5:**
14:	gb_est6:**
15:	em_estfun:**
16:	em_estom:**
17:	em_gss_hum:**
18:	em_gss_inv:**
19:	em_gss_plr:**
20:	em_gss_vrt:**
21:	em_gss_fun:**
22:	em_gss_man:**
23:	em_gss_mus:**
24:	em_gss_pro:**
25:	em_gss_rod:**
26:	em_gss_phg:**
27:	em_gss_vxl:**
28:	gb_gss1:**
29:	gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	781.4	68.4	1889 11 AK014524 Mus muscu
2	747.8	65.4	786 10 BG185954
C 3	604.4	52.9	638 9 AW381394
C 4	601	52.6	638 9 AW381390 PMO-HT030

5	559.4	48.9	1026	10	BG682538
6	536	46.9	824	12	BI871335
7	497.2	43.5	945	14	BY1713938
8	496.4	43.4	498	13	BY284040
9	451.2	39.5	906	14	CB590890
10	405	35.4	405	10	BG319597
11	384.4	33.6	627	10	BF477923
12	321.8	28.2	363	14	D82606
13	321.8	28.2	363	14	T48573
14	262.8	23.0	428	14	CB794396
15	234.2	20.5	1688	11	AK009003
16	234.2	20.5	1700	11	AK009855
17	213.2	18.7	1619	11	AK940697
18	213.2	18.7	1688	11	AK009018
19	211.6	18.5	1736	11	BC008483
20	209.6	18.3	1609	11	AK030670
21	208.8	18.3	1422	11	AK084785
22	207.2	18.1	1344	11	BC008399
23	204	17.8	1633	11	AK003220
24	196.6	17.2	1962	11	AK078753
25	196	17.1	381	13	BY139401
26	187.2	16.4	896	10	BG682440
27	186	16.3	186	12	BI335844
28	181	15.8	1289	11	AK003930
29	178.6	15.6	1662	11	AK003650
30	178.4	15.6	1289	11	AK008914
31	178.4	15.6	3468	11	AK036537
32	178	15.6	2940	11	AK036811
33	177.8	15.6	1289	11	AK076129
34	175.2	15.3	1719	11	BC011076
35	173.2	15.2	754	13	EX999033
36	173	15.1	1266	11	AK052580
37	168.2	14.7	2102	11	AK028382
38	164.4	14.4	1952	11	AK090049
39	164.4	14.4	1965	11	AK081487
40	161.8	14.2	1693	11	AK018226
41	161.4	14.1	1922	11	AK049429
42	160.4	14.0	2671	11	AK032659
43	159.6	14.0	858	10	BG199939
44	159.2	13.9	2260	11	AK035624
45	158	13.8	1886	11	AK004500

## ALIGNMENTS

RESULT 1  
AK014524  
LOCUS  
DEFINITION  
1889 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631416M05 product:serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 7, full insert sequence.  
ACCESSION  
VERSION AK014524.1 GI:12852434  
KEYWORDS HTC; CAP trapper.  
SOURCE  
ORGANISM Mus musculus  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99275253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, I., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159







```

Db      604 CAAATCTATGATGACCAATCCAGGCGATGACCTCTAAGTATGTTAGGTAATTT 663
QY      826 TTTCCTCAGTTCAAGATAGAGAAGATTATGAAATGAAACAATATTTCAGAGCCCTAGGG 885
Db      664 TTTCCTCAGTTCAAGATAGAGAAGATTATGAAATGAAACAATATTTCAGAGCCCTAAGG 723
QY      886 CTGAAGATATC-TTTCATGATCCAAAGCAGTCTCTCGGATTCCTTCGGGGGGTGC 944
Db      724 CTGAAGATATC-TTTCATGATCCAAAGCAGTCTCTCGGATTCCTTCGGGGGGTGC 783
QY      945 TCT 947
Db      784 TCT 786

RESULT 3
AW381394/c
LOCUS      PM0-HT0302-271099-001-e08 HT0302 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW381394
VERSION    AW381394.1 GI:6886053
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 638)
            HGCP http://www.ludwig.org.br/ORESTES.
            The FAPESP/LICR Human Cancer Genome Project
            Unpublished
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM0&t2=PM0-HT0302-271099-001-e08&t3=1999-10-27&t4=1)
            Seg primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 636.
FEATURES   source
            Location/Qualifiers
                i..638
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /dev_stage="Adult"
                /clone_lib="HT0302"
                /note="Organ: head neck; Vector: puc18; Site 1: Smar;
                Site 2: Smal; A mini-library was made by cloning products
                derived from ORESTES PCR (U.S. Letters Patent application
                No. 196,716 - Ludwig Institute for Cancer Research)
                profiles into the pUC 18 vector. Reverse transcription of
                tissue mRNA and cDNA amplification were performed under
                low stringency conditions."
            BASE COUNT  175 a 149 c 132 g 182 t
            ORIGIN

Query Match      52.9%; Score 604.4; DB 9; Length 638;
Best Local Similarity 99.8%; Pred. No. 3.7e-154;
Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      32 TTGCTTCAACTGTTTCAGAGAGATGATCAATCAAGGAATGGAATGTTCTTTT 91
Db      608 TTGCTTCAACTGTTTCAGAGAGATGATCAATCAAGGAATGGAATGTTCTTTT 549
QY      92 CCTCTCTGAGCCCTCTTCGTCGCCCTGGCCCTGGTCCTGGCCCTCAAGATGATCC 151
Db      548 CCTCTCTGAGCCCTCTTCGTCGCCCTGGCCCTGGTCCTGGCCCTCAAGATGATCC 489

```

```

QY      152 TCTCTCAGATTCATTAAGTTGCTTCATGTTTAACTGCTCAGGATATGAACTCTTTCTA 211
Db      488 TCTCTCAGATTCATTAAGTTGCTTCATGTTTAACTGCTCAGGATATGAACTCTTTCTA 429
QY      212 ATAGTCAGTCAGGGCTCCAGTCTCACTGAAAGAGTTTTTTTCTGATATAATCATCCC 271
Db      428 ATAGTCAGTCAGGGCTCCAGTCTCACTGAAAGAGTTTTTTTCTGATATAATCATCCC 369
QY      272 ACAAGATATTCATCTCAGCATTTGATATGGGCTTTTTTGTGAAAAAGTGTATGGCTTT 331
Db      368 ACAAGATATTCATCTCAGCATTTGATATGGGCTTTTTTGTGAAAAAGTGTATGGCTTT 309
QY      332 ATAGAGCTACATTTAGAGTGTGCGGAAAAAATATATACGATGCCAAAGTGGAGGAGTTC 391
Db      308 ATAGAGCTACATTTAGAGTGTGCGGAAAAAATATATACGATGCCAAAGTGGAGGAGTTC 249
QY      392 TTACGATCATTTAGAGAGCACTAGACGTAATATTAATTAAGTGGGTTGAAAAATGAACAC 451
Db      248 TTACGATCATTTAGAGAGCACTAGACGTAATATTAATTAAGTGGGTTGAAAAATGAACAC 489
QY      452 ATGCGAAAAATCAAGAACTGATTTGGTGAAGTGGCATAAGCTCATCTGCTGTATATGTTGC 511
Db      188 ATGCGAAAAATCAAGAACTGATTTGGTGAAGTGGCATAAGCTCATCTGCTGTATATGTTGC 129
QY      512 TGGTGAATGCTGTGTACTTCAAGGCAAGTGGCAATCAGCTTCCACCAAGAGCGAACA 571
Db      128 TGGTGAATGCTGTGTACTTCAAGGCAAGTGGCAATCAGCTTCCACCAAGAGCGAACA 69
QY      572 TAAATTGGCCATTTCAAAATCTCCCAAGTCTCTGGGAAGGCGAGTCGCATGATGCATCAGG 631
Db      68 TAAATTGGCCATTTCAAAATCTCCCAAGTCTCTGGGAAGGCGAGTCGCATGATGCATCAGG 9
QY      632 AACGGA 637
Db      8 AACGGA 3

RESULT 4
AW381390/c
LOCUS      PM0-HT0302-271099-001-d11 HT0302 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW381390
ACCESSION  AW381390
VERSION     AW381390.1 GI:6886049
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 638)
            HGCP http://www.ludwig.org.br/ORESTES.
            The FAPESP/LICR Human Cancer Genome Project
            Unpublished
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM0&t2=PM0-HT0302-271099-001-d11&t3=1999-10-27&t4=1)
            Seg primer: puc 18 forward
            High quality sequence start: 30
            High quality sequence stop: 562.
FEATURES   source
            Location/Qualifiers
                i..638
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"

```

```
/dev stage="Adult"
/clone lib="HMO302"
/notes="Organ: head neck; Vector: puc18; Site: 1; SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 174 a 152 c 131 g 181 t
ORIGIN
```

```
Query Match 52.6%; Score 601; DB 9; Length 638;
Best Local Similarity 98.3%; Pred. No. 3.1e-153;
Matches 618; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
```

```
QY 9 CCTTGTCTGCGAAGTGGAGAGTTTCTTCAACTGTTTCAGAGAGTGGATGCAATCA 68
DB 629 CCTTGGGCAACAAATGAGAGTTTCTTTCAGCCGTTTCAGAGAGTGGATGCAATCA 570
QY 69 AGGAATGGAATGTTCTTCTCTCTGAGCCTCTTCTGCTGCCCTGGCCCTGGTCG 128
DB 569 AGGAATGGAATGTTCTTCTCTCTGAGCCTCTTCTGCTGCCCTGGCCCTGGTCG 511
QY 129 CTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGTTCAATGTTAACTGC 188
DB 510 CTGGGCGCTCAAGATGACTCCCTCTCTCAGATTGATAGTTGTTCAATGTTAACTGC 451
QY 189 CTCAGGATGGAATCTTCTTAATAGTAGTCAGGCTCCAGTCTCAACTGAAAGAGT 248
DB 450 CTCAGGATGGAATCTTCTTAATAGTAGTCAGGCTCCAGTCTCAACTGAAAGAGT 391
QY 249 TTTTCTGATATAATGATCCACAGGATTTATGATCTCAGCATTTGATGGCTTTT 308
DB 390 TTTTCTGATATAATGATCCACAGGATTTATGATCTCAGCATTTGATGGCTTTT 331
QY 309 TGCTGAAAGATGTATGGCTTTTCATAGGACTACATTTGAGTGTCCGAAATATACGA 368
DB 330 TGCTGAAAGATGTATGGCTTTTCATAGGACTACATTTGAGTGTCCGAAATATACGA 271
QY 369 TGCCAAAGTGGAGCGATGTTGCTTACGAATCATTTTGAAGAAGACACTAGACCTAATA 428
DB 270 TGCCAAAGTGGAGCGATGTTGCTTACGAATCATTTTGAAGAAGACACTAGACCTAATA 211
QY 429 TAAGTGGTGAATGAATGAACACATGCGAAATCAAGAGCTGATTTGTTGAGTGGCAT 488
DB 210 TAAGTGGTGAATGAATGAACACATGCGAAATCAAGTGGATTTGTTGAGTGGCAT 151
QY 489 AAGTCATCTGCTGTAATGTTGCTGGTGAATGCTGTACTTCAAGGCAAGTGGCAATC 548
DB 150 AAGTCATCTGCTGTAATGTTGCTGGTGAATGCTGTACTTCAAGGCAAGTGGCAATC 91
QY 549 AGCCTTCCACAGAGCGAAGCAATTAATGGCATTTCAATCTCCCAAGTCTCTGGAA 608
DB 90 AGCCTTCCACAGAGCGAAGCAATTAATGGCATTTCAATCTCCCAAGTCTCTGGAA 31
QY 609 GGCAGTCGCCATGATGATCAGGAACGGA 637
DB 30 GGCAGTCGCCATGATGATCAGGAACGGA 2
```

```
RESULT 5
BG682538
LOCUS 602624811F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749699 5',
DEFINITION mRNA sequence.
ACCESSION BG682538
VERSION BG682538.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

# REFERENCE

1 (bases 1 to 1026)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: James Cleaver, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Place: LLNL0603 row: f column: 04  
High quality sequence start: 3  
High quality sequence stop: 791.  
Location/Qualifiers

# FEATURES

1..1026  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4749699"  
/tissue\_type="squamous cell carcinoma"  
/lab\_host="NCI CGAP Skn4"  
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.5kb. Library constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 303 a 221 c 213 g 289 t  
ORIGIN

Query Match 48.9%; Score 559.4; DB 10; Length 1026;  
Best Local Similarity 99.8%; Pred. No. 9.9e-142;  
Matches 560; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 583 TTCAAATCTCCAAAGTGTCTTGGAGGCGAGTCGCGCATGATCAGGAACGGAAGTTC 642
DB 9 TTCAAATCTCCAAAGTGTCTTGGAGGCGAGTCGCGCATGATCAGGAACGGAAGTTC 68
QY 643 AATTGTCTGTATTTGAGGAGCCCATCAATGAGATCTTGGAGCTCAGATACATGGTGGC 702
DB 69 AATTGTCTGTATTTGAGGAGCCCATCAATGAGATCTTGGAGCTCAGATACATGGTGGC 128
QY 703 ATAAACATGTACCTTCTCTCTCCCTGAGAAATGACCTCTCTGAAATTTGAAACCAACTGACC 762
DB 129 ATAAACATGTACCTTCTCTCTCCCTGAGAAATGACCTCTCTGAAATTTGAAACCAACTGACC 188
QY 763 TTTCCAGATCTAATGGAATGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTA 822
DB 189 TTTCCAGATCTAATGGAATGACCAATCCAGGCGAATGACCTCTAAGTATGTTGAGTA 248
QY 823 TTTTCTCTCAGTCAAGATAGAGAAATATGAAATGAAACAAATTTTGAGAGCCCTA 882
DB 249 TTTTCTCTCAGTCAAGATAGAGAAATATGAAATGAAACAAATTTTGAGAGCCCTA 308
QY 883 GGGCTGAAAGATATCTTTGATGAAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGT 942
DB 309 GGGCTGAAAGATATCTTTGATGAAATCCAAAGCAGATCTCTCTGGGATTCCTTCGGGGGGT 368
QY 943 GGTCTGTATATATCAGGATGATGCAATTTTACATAGAGTCACTGAGAGGGGCAAC 1002
DB 369 GGTCTGTATATATCAGGATGATGCAATTTTACATAGAGTCACTGAGAGGGGCAAC 428
QY 1003 GAGGTACTGCTGCCACAGGAAGTAAATTTGAGAAAGCAACTCCCTCAGTCCAGCGCTG 1062
DB 429 GAGGTACTGCTGCCACAGGAAGTAAATTTGAGAAAGCAACTCCCTCAGTCCAGCGCTG 488
QY 1063 TTTAGAGCTGACCAACCATTCCTTATTTGTTATCAGGAAGGATGACATCATCTATTTCAGT 1122
DB 489 TTTAGAGCTGACCAACCATTCCTTATTTGTTATCAGGAAGGATGACATCATCTATTTCAGT 548
QY 1123 GGCAGAGTTTCTTGGCCCTTGA 1143
```

```

Db      549  GGCAAGATCTCTCGGAGTCTCGGGGGTCTCTGTATATATCAAGGATGTC 371
|||||
1088  ATATTGTAGAAAAGCAACTCCCTCAGTCACGCTGTTTAGAGCTGACCAACCATCTCTAT 1087
Db      432  ATATTGTAGAAAAGCAACTCCCTCAGTCACGCTGTTTAGAGCTGACCAACCATCTCTAT 491
|||||
1088  TTGTTATCAGGAGGATGACATCATCTTATTCAGTGCACAGTTCTTGCCTTGA 1143
Db      492  TTGTTATCAGGAGGATGACATCATCTTATTCAGTGCACAGTTCTTGCCTTGA 547
|||||

RESULT 7
BY713938
LOCUS      945 bp      mRNA      linear      EST 17-DEC-2002
DEFINITION RIKEN full-length enriched, 0 day neonate skin Mus
VERSION     BY713938
KEYWORDS    Mus musculus (house mouse)
SOURCE      EST
ORGANISM    Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
Okazaki.Y., Furuno.M., Kasukawa.T., Adachi.J., Bono.H., Kondo.S.,
Nikaido.I., Osato.N., Saito.R., Suzuki.H., Yamanaka.I., Kiyosawa.H.,
Yagi.K., Tomaru.Y., Hasegawa.Y., Nogami.A., Schonbach.C.,
Gojobori.T., Baldarelli.R., Hill.D.P., Bult.C., Hume.D.A.,
Quackenbush.J., Schriml.E.M., Kanapin.A., Matsuda.H., Batalov.S.,
Beisel.X.W., Blake.J.A., Bradt.D., Brusic.V., ChoChia.C., Corbani
L.E., Cousins.S., Dalla.E., Dragani.T.A., Fletcher.C.F., Forrest
A., Frazer.K.S., Gaasterland.T., Gariboldi.M., Gissi.C., Godzik.A.,
Gough.J., Grimmond.S., Gustincich.S., Hirokawa.N., Jackson.I.J.,
Jarvis.E.D., Kanaya.A., Kawaji.H., Kawasawa.Y., Kedzierski.R.M.,
King.B.L., Konagaya.A., Kurochkin.I.V., Lee.Y., Lenhard.B., Lyons
P.A., Maglott.D.R., Maltai.L., Marchionni.L., McKenzie.L., Miki
H., Nagashima.T., Numata.K., Okido.T., Pavan.W.J., Perle.G.,
Pesole.G., Petrovsky.N., Pillai.R., Pontius.J.U., Qi.D.,
Raschadran.S., Ravasi.T., Reed.J.C., Reed.D.J., Reid.J., Ring
B.Z., Ringwald.M., Sandelin.A., Schneider.C., Sempel.C.A., Setou
M., Shimada.K., Sultana.R., Takenaka.Y., Taylor.M.S., Teasdale
R.D., Tomita.M., Verardo.R., Wagner.L., Wahlestedt.C., Wang.Y.,
Watanabe.Y., Wells.C., Wilming.L.G., Wynshaw-Boris.A., Yanagisawa
M., Yang.I., Yang.L., Yuan.Z., Zavolan.M., Zhu.Y., Zimmer.A.,
Carninci.P., Hayatsu.N., Hirozane-Kishikawa.T., Konno.H., Nakamura
M., Sakazume.N., Sato.K., Shiraki.T., Waki.K., Kawai.J., Aizawa.K.,
Arakawa.T., Fukuda.S., Hara.A., Hashizume.W., Imotani.K., Ishii
Y., Itoh.M., Kagawa.I., Miyazaki.A., Sakai.K., Sasaki.D., Shibata
K., Shinagawa.A., Yasunishi.A., Yoshino.M., Waterston.R., Lander
E.S., Rogers.J., Birney.E. and Hayashizaki.Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED
12466851
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp
URL: http://genome.gsc.riken.go.jp/
Adachi.O., Aizawa.K., Akimura.T., Arakawa.T., Carninci.P., Fukuda
S., Hashizume.W., Hayashida.K., Hirozane.T., Hori.F., Imotani.K.,
Ishii.Y., Itoh.M., Kagawa.I., Kawai.J., Kojima.Y., Kondo.S., Konno
H., Koya.S., Miyazaki.A., Murata.M., Nakamura.M., Nomura.K.,

```

```

Db      549  GGCAAGATCTCTCGGAGTCTCGGGGGTCTCTGTATATATCAAGGATGTC 371
|||||
1088  ATATTGTAGAAAAGCAACTCCCTCAGTCACGCTGTTTAGAGCTGACCAACCATCTCTAT 1087
Db      432  ATATTGTAGAAAAGCAACTCCCTCAGTCACGCTGTTTAGAGCTGACCAACCATCTCTAT 491
|||||
1088  TTGTTATCAGGAGGATGACATCATCTTATTCAGTGCACAGTTCTTGCCTTGA 1143
Db      492  TTGTTATCAGGAGGATGACATCATCTTATTCAGTGCACAGTTCTTGCCTTGA 547
|||||

RESULT 6
BI871335
LOCUS      824 bp      mRNA      linear      EST 11-OCT-2001
DEFINITION 603392761F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402566 5',
mRNA sequence.
ACCESSION  BI871335
VERSION     BI871335
KEYWORDS    Homo sapiens (human)
SOURCE      EST
ORGANISM    Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 824)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: rsb@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2027 row: 1 column: 15
High quality sequence stop: 824.
Location/Qualifiers
1..824
/organism="Homo sapiens"
/molecule_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5402566"
/tissue types="adenocarcinoma, cell line"
/lab host="NIH_MGC_90"
/clone lib="NIH_MGC"
/Note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT      232 a 186 c 164 g 242 t
ORIGIN
Query Match      46.9%; Score 536; DB 12; Length 824;
Best Local Similarity 100.0%; Pred. No. 2.3e-135;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      608  AGCAGTCCCATGATGATCAGCAAGGAGTTCATTTGCTGTTTATTCAGGACCAT 667
Db      12  AGGATGTCGCGATGATGATCAGCAAGGAGTTCATTTGCTGTTTATTCAGGACCAT 71
|||||
668  CAATGAAGATCTTGTGAGTCAGATACATGTTGGTGCATAAACAATGATCTGTCGCTG 727
Db      72  CAATGAGATCTTGTGAGTCAGATACATGTTGGTGCATAAACAATGATCTGTCGCTG 131
|||||
728  AGAATGACCTCTGTAATTAAGAAACAACATGACCTTTCAGAACTTAATGAATGACCA 787
Db      132  AGAATGACCTCTGTAATTAAGAAACAACATGACCTTTCAGAACTTAATGAATGACCA 191
|||||
788  ATCCAGGCGGATGACCTCTAAGTATGTTGAGGTATTTTCTCAGTCAAGATGAGA 847
Db      192  ATCCAGGCGGATGACCTCTAAGTATGTTGAGGTATTTTCTCAGTCAAGATGAGA 251
|||||
848  AGAATTATGAATGAACAAATATTTGAGAGCCCTAGGCTGGAAGATATCTTTGATGAT 907
Db      252  AGAATTATGAATGAACAAATATTTGAGAGCCCTAGGCTGGAAGATATCTTTGATGAT 311
|||||
908  CCAAGACAGATCTCTCTGGGATTTGCTTCGGGGGGTCTCTGTATATATCAAGGATGATGC 967

```



source

1...498  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:9980410603 ; IMAGE:4749699"  
 /tissue\_type="squamous cell carcinoma"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Skn4"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;  
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.5kb. Library constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 153 a 101 c 114 g 130 t  
 ORIGIN

Query Match 43.4%; Score 456.4; DB 13; Length 498;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-124;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 593 TTCAATCTCCCAAGTCTCTGGGAAGCAGTCCCATGATGATCAGGACGGAAGTTC 642  
 Db 1 TTCAATCTCCCAAGTCTCTGGGAAGCAGTCCCATGATGATCAGGACGGAAGTTC 60  
 QY 643 AATTCTCTGTATTAGGACCCATCAATGAAGATCTTGGCTCAGATCAATGTGGC 702  
 Db 61 AATTCTCTGTATTAGGACCCATCAATGAAGATCTTGGCTCAGATCAATGTGGC 120  
 QY 703 ATAAACATGAGTCTCTGGCTGAGATGACCTCTGAAATGAAACAAACTGACC 762  
 Db 121 ATAAACATGAGTCTCTGGCTGAGATGACCTCTGAAATGAAACAAACTGACC 180  
 QY 763 TTTCAGAACTAATGAATGACCATCCAGGCGATGACCTTAAGTATGTAGSTA 822  
 Db 181 TTTCAGAACTAATGAATGACCATCCAGGCGATGACCTTAAGTATGTAGSTA 240  
 QY 923 TTTTTCCTCAGTTCAGATGAGAAATATGAATGAACAAATATTTGAGAGCCCTA 882  
 Db 241 TTTTTCCTCAGTTCAGATGAGAAATATGAATGAACAAATATTTGAGAGCCCTA 300  
 QY 883 GGCTGAAGATATCTTGAATGATCCAAAGCAGATCTCTGGGAGTCTGGGGGT 942  
 Db 301 GGCTGAAGATATCTTGAATGATCCAAAGCAGATCTCTGGGAGTCTGGGGGT 360  
 QY 943 CGTCTGATATACAGATGATGACAAATCTTACATAGAGTCACTGAGGAGGCACC 1002  
 Db 361 CGTCTGATATACAGATGATGACAAATCTTACATAGAGTCACTGAGGAGGCACC 420  
 QY 1003 GAGGTACTGTGCCACAGGAAGTAAATTTGAGAAAGCAACTCCCTCAGTCCAGCTG 1052  
 Db 421 GAGGTACTGTGCCACAGGAAGTAAATTTGAGAAAGCAACTCCCTCAGTCCAGCTG 480  
 QY 1063 TTTAGAGTGCACCCCA 1080  
 Db 481 TTTAGAGTGCACCCCA 498

RESULT 9

CB590890

LOCUS

DEFINITION AGENCOURT\_12567930 NIH\_MGC\_136 Mus musculus CDNA clone  
 IMAGE:30288309 5', mRNA sequence.

ACCESSION

CB590890

VERSION

CB590890.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 906)

AUTHORS

NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. David Rowe

CDNA Library Preparation: Invitrogen Corp

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Cloned through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NDAM325 row: i column: 22

High quality sequence stop: 668.

Location/Qualifiers

1...906

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:30288309"

/tissue\_type="embryonic limb, maxilla and mandible"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_136"

/note="Vector: pCMV-SPORT6.1.cdb; Site:1: EcoRV; Site:2:

NotI; Normalized, full-length enriched library from pool

of mouse embryonic limb, maxilla and mandible, embryonic

day 17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and

jaw equivalents from respective days). Cloned

directionally, oligo-dT primed

(5'-GACTAGTCTAGATCGAGCGCCGCC (T)15-3'. Size selected

for the &gt;1kb fragments, average insert size 1.2 kb.

Normalization to Cot 7.5. Tissue contributed by David

Rowe; library constructed by ResGen, Invitrogen Corp.

Note: this is a NIH\_MGC Library."

BASE COUNT 285 a 186 c 204 g 231 t

ORIGIN

Query Match 39.5%; Score 451.2; DB 14; Length 906;

Best Local Similarity 79.5%; Pred. No. 3.9e-112;

Matches 534; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTCTGCGAAGATGCAAGATTGCTTCAACTTCAGAGATGGAT 60

Db 54 ATGGCTCCCTTCTGCGAAGATGCAAGATTGCTTCAACTTCAGAGATGGAT 113

QY 61 GACAACTAGGAAATGCAAGATTGCTTCAACTTCAGAGATGGAT 120

Db 114 AGTAGCAGGAAATGCAAGATTGCTTCAACTTCAGAGATGGAT 173

QY 121 CTGGTCCGCTTGGGCGCTCAAGATGCTCCCTCTCTCAGATTGATAAGTTCATGTT 180

Db 174 CTAACTCGTCTGGTGTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 233

QY 181 AACACTGCTCAGGATGAAACTCTCTTAATAGTCTAGGCTCAGGCTCAGTCTCACTG 240

Db 234 AACATACCATCAAGCAAGGAAACTCTTAATAGTCTAGGCTCAGGCTCAGTCTCACTG 293

QY 241 AAAAGAGTCTTCTGATATAAATGATCCCAAGGATTATGATCTCAGATTGTGAAT 300

Db 294 AAAAGAGTCTTCTGATATAAATGATCCCAAGGATTATGATCTCAGATTGTGAAT 353

QY 301 GGGCTTTTCTGATAAAGTGTATGCTTTTCAAGGATCTAATGAGTGTGCGGAAAAA 360

Db 354 GAGTCTTTTGGCAGAAAAAGTGTATGCTTTTCAAGGATCTAATGAGTGTGCGGAAAAA 413

QY 361 TTATACGATGCAAGTGGAGGCTTCACTTTTACGAACTCATTAGACACTAGACCT 420

Db 414 TTATACGATGCAAGTGGAGGCTTCACTTTTACGAACTCATTAGACACTAGACCT 473

QY 421 AATATTAAATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480

Db 474 AATATTAAATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 533

QY 481 GGTGGGCAATAGCTCATCTGCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540

Db 534 AGCAGGCTCAGCTCGGCTGT 593



## ORIGIN

Query Match	28.2%;	Score 321.8;	DB 14;	Length 363;
Best Local Similarity	93.7%;	Pred. No. 6.7e-77;		
Matches 340;	Conservative 0;	Mismatches 22;	Indels 1;	Gaps 1;
<hr/>				
QY	492	CTCATCTGCTGTAATGGFHGCTGCTGAATGCTGTACTTCAAAGGCGAAGTGGCAATCAGC	551	
DB	1	CTCATCTGCTGTAATGGTGTGTGTGATGCTGTACTTCAAAGGCGAAGTGGNAATCAGC	60	
QY	552	TTTCAACAAGAGCGAAACCAATAATTGGCCATTTCAAATCTCCCAAGTGCCTTGGGAAGCG	611	
DB	61	CTTCAACCAAGAGCGAAACCAATAATTGGCCATTTCAAATCTCCCAAGTGCCTTGGGAAGCG	120	
QY	612	AGTCGGCATGATGCATCAGGAACGGGAAGTTCAAATTTGTCTGTATTGAGGACCCCATCAAT	671	
DB	121	AGCGCCGCGATGCATCAGGAACGGGAGTCACTTTGTCTGTATTGAGGACCCATCANT	180	
QY	672	GAAGATCTTTGAGCTCAGATACAATGGTGGCATAAACAATGTACGTTCTGCTGCCTGAGAA	731	
DB	181	GAAGATCTTTGAGCTCAGATACAATGGTGGCATAAACAATGTACGTTCTGCTGCCTGAGAA	240	
QY	732	TGACCTCTCTCGAATTTGAAACCAACTGACCTTTTCAGAACTCTAATGGAAATGGACCAATCC	791	
DB	241	TGACCTCTCTCGAATTTGAAACCAACTGACCTTTTCAGAACTCTAATGGAAATGGACCAATCC	300	
QY	792	AAGCGGAATGACCTCTAAGTATGTTGAGGTATTTT-TTCCTCAGTTCAAGATAGAGAAGA	850	
DB	301	ACGGGGGAATWACCTCTAAGTATGTNGGTATTTTTCCTCCCAAGCTCANGATTTGAGAAGG	360	
QY	851	ATT 853		
DB	361	AAT 363		
<hr/>				
RESULT 13				
T48573				

## DEFINITION

7 end; MAFNA sequence.  
 T48573  
 T48573.1 GI:642773  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 363)  
 Bell, G.I. and Takeda, J.  
 Human pancreatic islet cDNAs  
 Unpublished  
 Contact: Bell GI or Takeda J  
 HHMI  
 Univ. of Chicago  
 5841 S. Maryland Ave., MC1028, Chicago IL 60637  
 Tel: 3127039116  
 Fax: 3127030271  
 Email: g-bell@uchicago.edu  
 Seq primer: SK primer.  
 Location/Qualifiers  
 1..363  
 source

```

1_c7b="maca"
/db_xref="taxon:9606"
/clone="hbc2926"
/clone_lib="Human pancreatic islet"
/clone="Vector; Lambda ZAP1; Site 1: Eco RI; Site 2: Xho
I; mRNA was prepared from normal adult human islets. cDNA
was directionally synthesized from the Xho I in the vector
to the EcoRI site. cDNA was size fractionated to remove
sequences <1000 bp in size."
BASE COUNT: 104 a 78 c 82 g 89 t 10 others
ORIGIN

```

Query Match 28.2%; Score 321.8; DB 14; Length 363;  
 Best Local Similarity 93.7%; Pred. No. 6.7e-77;  
 Matches 340; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 492 CTCATCTGCTGAATGGTCTGCTGAATGCTGTGACTTCAAAAGGCAAGTGGCAATCAGC 551  
 DB 1 CTCATCTGCTGAATGGTGGTGAATGCTGTGACTTCAAAAGGCAAGTGGCAATCAGC 60

QY 552 CTTCCACAGAGGCAAAACATTAATTCGCAATTCCTCAATTCCTCAATTCCTGGGAAGGC 611  
 DB 61 CTTCCACAGAGGCAAAACATTAATTCGCAATTCCTCAATTCCTCAATTCCTGGGAAGGC 120

QY 612 AGTCGCCATGATCATCAGGAACGGAAGTTCATTTGCTTATTCAGGACCCATCAT 671  
 DB 121 AGCCGCCANGTGCATCAGGAACGGAAGTTCATTTGCTTATTCAGGACCCATCAT 180

QY 672 GAAGATCTTGAGCTCAGATACATAGTGGTCAATAACATGATGCTGCTGCTGAGAA 731  
 DB 181 GAAGATCTTGAGCTCAGATACATAGTGGTCAATAACATGATGCTGCTGCTGAGAA 240

QY 732 TGACCTCTGGAATGGAACAACTGACCTTTCAGATCTTAATGGATGGACCAATCC 791  
 DB 241 TGACCTCTGGAATGGAACAACTGACCTTTCAGATCTTAATGGATGGACCAATCC 300

QY 792 AAGCGCAATGACCTCTAAGTATGTTGAGTATTTT-TTCCTCAGTTTCAAGATAGAGAA 850  
 DB 301 ACGGGAATNACTCTAAGTATGTTGAGTATTTTTCCTCAGTTTCAAGATAGAGAA 950

QY 851 ATT 853  
 DB 361 AAT 363

RESULT 14  
 LOCUS CB794396  
 DEFINITION AMGNNUC:SRPB2-00314-A2-A srpb2 (10220) Rattus norvegicus cDNA clone  
 srpb2-00314-a2 5', mRNA sequence.  
 ACCESSION CB794396  
 VERSION CB794396.1 GI:29882873  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE  
 1 (bases 1 to 428)  
 Angen EST Program.  
 TITLE Angen Rat EST Program  
 JOURNAL Unpublished  
 COMMENT Contact: Dan Fitzpatrick  
 Angen, Inc  
 One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA  
 Tel: 805 447-4881  
 Plate: 00314 row: a column: 2.  
 Location/Qualifiers  
 1..428  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="srpb2-00314-a2"  
 /tissue\_type="prostate tissue"  
 /clone\_id="srpb2 (10220)"  
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; rat  
 prostate normalized double selected poly(A+), mRNA size  
 Fraction > 1 kb"

FEATURES  
 source  
 126 a 88 c 93 g 121 t

Query Match 23.0%; Score 262.9; DB 14; Length 428;  
 Best Local Similarity 78.4%; Pred. No. 1e-60;  
 Matches 315; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 1 ATGGCTCCCTTGTCTGACGAATGCGAGTGTTCCTCAACCTGTTCAGAGAGATGGAT 60  
 DB 26 ATGGCTCCCTTGTCTGACGAATGCGAGTGTTCCTCAACCTGTTCAGAGAGATGGAT 85

QY 61 GACATCAGGAATGGAATGTGTTCTTCTCTCTGAGCCTCTTCGCTGCTGCTGCTGCTG 120  
 DB 86 AGTAGTCAGGAATGGAATGTGTTCTTCTCTCTGAGCCTCTTCGCTGCTGCTGCTGCTG 145

QY 121 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 DB 146 CTAATCCGTTGGTCTGCTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 205

QY 181 AACACTGCTCAGGATATGGAATCTCTTCTAATAGTCAAGTCTGAGGTCTCAAGTCTCA 240  
 DB 206 ATCTCCCATCAAGCAAGGAAATTCATCGACAGTCAAGTCTGAGGTCAATATCAATG 265

QY 241 AAAGAGTTTTTCTCATATTAATGATCCCAAGGATTAATGATCTCAGCAATTTGTAAT 300  
 DB 266 AAAGAGTTTTTCTCATATAAATCTCATCTCAAGGATTAATGATCTCAGCAATTTGCA 325

QY 301 GGGCTTTTCTGAAAAAGTGTATGGCTTTCATTAAGGACTACATTCAGTGTGCTGCTG 360  
 DB 326 GAGTTTTTCTGAGGAAGTATTTGATTTTCATTAAGGACTATATGAGTGTGCTGCTG 385

QY 361 TTATACGATCCCAAGTGGAGGAGTGTGATTTTACGATCAT 402  
 DB 386 TTATACGATCCCAAGTGGAGGAGTGTGATTTTACGATCAT 427

RESULT 15  
 LOCUS AK009003  
 DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched  
 library, clone:2300002C09 product:weakly similar to SERPIN B11  
 [Homo sapiens], full insert sequence.

ACCESSION AK009003  
 VERSION AK009003.1 GI:12843530  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 92279253  
 PUBMED 10349636

REFERENCE  
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

REFERENCE  
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multiplexed cDNA libraries  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

REFERENCE  
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Harada, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,







Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	380	100.0	380	2	US-08-472-559-34	Sequence 34, Appl
2	380	100.0	380	2	US-08-474-661-34	Sequence 34, Appl
3	380	100.0	380	2	US-08-611-977-34	Sequence 34, Appl
4	19	5.0	19	2	US-08-611-977-3	Sequence 3, Appl
5	18	4.7	20	2	US-08-611-977-6	Sequence 6, Appl
6	17	4.5	17	2	US-08-611-977-4	Sequence 4, Appl
7	17	4.5	17	2	US-08-472-559-6	Sequence 6, Appl
8	17	4.5	20	2	US-08-474-661-6	Sequence 6, Appl
9	13	3.4	382	1	US-07-768-286B-6	Sequence 6, Appl
10	13	3.4	382	1	US-08-487-823B-3	Sequence 3, Appl
11	13	3.4	382	2	US-08-997-040-3	Sequence 3, Appl
12	13	3.4	382	2	US-09-203-237-3	Sequence 3, Appl
13	13	3.4	392	1	US-07-768-286B-4	Sequence 3, Appl
14	13	3.4	405	1	US-08-121-714-7	Sequence 4, Appl
15	13	3.4	405	1	US-08-477-108A-7	Sequence 7, Appl
16	13	3.4	405	2	US-08-477-112-7	Sequence 7, Appl
17	13	3.4	405	5	PCR-US93-083322-7	Sequence 7, Appl
18	13	3.4	415	1	US-07-911-531-19	Sequence 19, Appl
19	13	3.4	415	1	US-07-633-636A-19	Sequence 19, Appl
20	13	3.4	415	1	US-07-768-286B-2	Sequence 2, Appl
21	13	3.4	415	4	US-08-026-408-4	Sequence 4, Appl
22	13	3.4	434	1	US-07-679-052A-15	Sequence 15, Appl
23	13	3.4	438	1	US-07-679-052A-17	Sequence 17, Appl
24	11	2.9	17	2	US-08-472-659-4	Sequence 4, Appl
25	11	2.9	17	2	US-08-474-661-4	Sequence 4, Appl
26	10	2.6	10	2	US-08-611-977-9	Sequence 9, Appl
27	10	2.6	19	2	US-08-472-659-3	Sequence 3, Appl

;  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-472-659-34

Query Match 100.0%; Score 380; DB 2; Length 380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAGQDDSLSQIDKLLHV 60  
Db 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAGQDDSLSQIDKLLHV 60  
QY 61 NTASGYNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120  
Db 61 NTASGYNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120  
QY 121 LYDAKVERVDFTNHLEDFRRNINKWVENETHGKIKNVIGGGISSAVVWLVNAVYFKG 180  
Db 121 LYDAKVERVDFTNHLEDFRRNINKWVENETHGKIKNVIGGGISSAVVWLVNAVYFKG 180  
QY 181 WQSAFTKSETINCHFKSPKSGKAVAMHQERKNLSVIEDPSMKILRLYNGGINMYVL 240  
Db 181 WQSAFTKSETINCHFKSPKSGKAVAMHQERKNLSVIEDPSMKILRLYNGGINMYVL 240  
QY 241 LPENDLSEIENKLTQNLMEWTPREMTSKYVEVFPQFKIEKNYEMQYLRALGLKIDIF 300  
Db 241 LPENDLSEIENKLTQNLMEWTPREMTSKYVEVFPQFKIEKNYEMQYLRALGLKIDIF 300  
QY 301 DESKADLSGIASGGRLYISRMHKSIVTEEGTEATAATGSNIVEKQLPOSTLFRADHP 360  
Db 301 DESKADLSGIASGGRLYISRMHKSIVTEEGTEATAATGSNIVEKQLPOSTLFRADHP 360  
QY 361 FLFVIRKDDIILFSGKVSCP 380  
Db 361 FLFVIRKDDIILFSGKVSCP 380

## RESULT 2

US-08-474-661-34  
; Sequence 34, Application US/08474661  
; Patent No. 5874253  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Ruyuki  
; APPLICANT: TSURUOKA, No. 5874253uo  
; APPLICANT: NAKAZATO, Hiroshi  
; APPLICANT: MIURA, Kenju  
; APPLICANT: ISHIDA, No. 5874253uhiro  
; APPLICANT: KURIHARA, Tatsuya  
; APPLICANT: YAMAICHI, Kozo  
; APPLICANT: YAMAGUCHI, No. 5874253cmi  
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: George Mason Bldg., Washington & Prince Sts.  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474,661  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/091,028

;  
; FILING DATE: 14-JUL-1993  
; APPLICATION NUMBER: JP 4-212305  
; FILING DATE: 17-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-067339  
; FILING DATE: 04-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REA, TERESA STANEX  
; REGISTRATION NUMBER: 30,427  
; REFERENCE/DOCKET NUMBER: 001560-204  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-6620  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-474-661-34

Query Match 100.0%; Score 380; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAGQDDSLSQIDKLLHV 60  
Db 1 MASLAANAEEFCNLFREMDNQNGNVFFSSLSLFAALALVRLGAGQDDSLSQIDKLLHV 60  
QY 61 NTASGYNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120  
Db 61 NTASGYNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120  
QY 121 LYDAKVERVDFTNHLEDFRRNINKWVENETHGKIKNVIGGGISSAVVWLVNAVYFKG 180  
Db 121 LYDAKVERVDFTNHLEDFRRNINKWVENETHGKIKNVIGGGISSAVVWLVNAVYFKG 180  
QY 181 WQSAFTKSETINCHFKSPKSGKAVAMHQERKNLSVIEDPSMKILRLYNGGINMYVL 240  
Db 181 WQSAFTKSETINCHFKSPKSGKAVAMHQERKNLSVIEDPSMKILRLYNGGINMYVL 240  
QY 241 LPENDLSEIENKLTQNLMEWTPREMTSKYVEVFPQFKIEKNYEMQYLRALGLKIDIF 300  
Db 241 LPENDLSEIENKLTQNLMEWTPREMTSKYVEVFPQFKIEKNYEMQYLRALGLKIDIF 300  
QY 301 DESKADLSGIASGGRLYISRMHKSIVTEEGTEATAATGSNIVEKQLPOSTLFRADHP 360  
Db 301 DESKADLSGIASGGRLYISRMHKSIVTEEGTEATAATGSNIVEKQLPOSTLFRADHP 360  
QY 361 FLFVIRKDDIILFSGKVSCP 380  
Db 361 FLFVIRKDDIILFSGKVSCP 380

## RESULT 3

US-08-611-977-34  
; Sequence 34, Application US/08611977  
; Patent No. 5972886  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Ruyuki  
; APPLICANT: TSURUOKA, No. 5972886uo  
; APPLICANT: NAKAZATO, Hiroshi  
; APPLICANT: MIURA, Kenju  
; APPLICANT: ISHIDA, No. 5972886uhiro  
; APPLICANT: KURIHARA, Tatsuya  
; APPLICANT: YAMAICHI, Kozo  
; APPLICANT: YAMAGUCHI, No. 5972886cmi  
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404

RESULT 4  
US-08-611-977-3  
; Sequence 3, Application US/08611977

APPLICANT: ISHIDA, No. 5972886uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAUCHI, Kozo  
APPLICANT: YAMAGUCHI, No. 5972886omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/611,977  
FILING DATE: 06-MAR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-204  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
US-08-611-977-6

Query Match 4.7%; Score 18; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.5e-10;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGRLYISRM 322  
Db 1 ADLSGIASGRLYISRM 18

RESULT 6  
US-08-611-977-4  
Sequence 4, Application US/08611977  
Patent No. 5972886  
GENERAL INFORMATION:  
APPLICANT: TSUTIMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSURUOKA, No. 5972886uo  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. 5972886uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAUCHI, Kozo  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/611,977  
FILING DATE: 06-MAR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-204  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-611-977-4

Query Match 4.5%; Score 17; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 SYIEVTEGTEATAATG 341  
Db 1 SYIEVTEGTEATAATG 17

RESULT 7  
US-08-472-659-6  
Sequence 6, Application US/08472659  
Patent No. 5831030  
GENERAL INFORMATION:  
APPLICANT: TSUTIMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSURUOKA, No. 5831030uo  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. 5831030uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAUCHI, Kozo  
APPLICANT: YAMAGUCHI, No. 5831030omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,659
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-472-659-6

```

```

Query Match 4.5%; Score 17; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGGRLYISRM 321
Db 1 ADLSGIASGGRLYISRM 17

```

```

RESULT 8
US-08-474-661-6
; Sequence 6, Application US/08474661
; Patent No. 5874253
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, No. 5874253uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5874253uhiro
; APPLICANT: KURIHARA, Tatsuuya
; APPLICANT: YAWAICHI, Kozo
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swacker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,661
; FILING DATE: 07-JUN-1995

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: REA, TERESA STANEK
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-474-661-6

Query Match 4.5%; Score 17; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGGRLYISRM 321
Db 1 ADLSGIASGGRLYISRM 17

RESULT 9
US-07-768-286B-6
; Sequence 6, Application US/07768286B
; Patent No. 544153
; GENERAL INFORMATION:
; APPLICANT: GOSS, Neil H.
; APPLICANT: RICHARDSON, Michael A.
; TITLE OF INVENTION: VARIANTS OF PAI-2
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/768,286B
; FILING DATE: 19911011
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00603
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/157 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids

```

;  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: procein  
US-07-768-286B-6

Query Match 3.4%; Score 13; DB 1; Length 382;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181  
| | | | | | | | | |  
Db 163 MVLNNAVYFKGW 175

RESULT 10  
US-08-487-823B-3  
; Sequence 3, Application US/08487823B  
; Patent No. 5700924  
; GENERAL INFORMATION:  
; APPLICANT: Braxton, Scott M.  
; APPLICANT: Diep, Dinh  
; APPLICANT: Stuart, Susan G.  
; TITLE OF INVENTION: NOVEL SERPIN DERIVED FROM HUMAN  
; TITLE OF INVENTION: HYPOTHALAMUS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,823B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Luther, Barbara J.  
; REGISTRATION NUMBER: 33,954  
; REFERENCE/DOCKET NUMBER: PF-0039 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-852-0195  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 382 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
US-08-487-823B-3

Query Match 3.4%; Score 13; DB 1; Length 382;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181  
| | | | | | | | | |  
Db 163 MVLNNAVYFKGW 175

RESULT 11  
US-08-997-040-3  
; Sequence 3, Application US/08997040  
; Patent No. 5929210  
; GENERAL INFORMATION:  
; APPLICANT: Braxton, Scott M.  
; APPLICANT: Diep, Dinh  
; APPLICANT: Stuart, Susan G.  
; TITLE OF INVENTION: NOVEL SERPIN DERIVED FROM HUMAN  
; TITLE OF INVENTION: HYPOTHALAMUS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/997,040  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,823  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Luther, Barbara J.  
; REGISTRATION NUMBER: 33,954  
; REFERENCE/DOCKET NUMBER: PF-0039 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-852-0195  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 382 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
US-08-997-040-3

Query Match 3.4%; Score 13; DB 2; Length 382;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181  
| | | | | | | | | |  
Db 163 MVLNNAVYFKGW 175

RESULT 12  
US-09-203-237-3  
; Sequence 3, Application US/09203237  
; Patent No. 5955284  
; GENERAL INFORMATION:  
; APPLICANT: Braxton, Scott M.  
; APPLICANT: Diep, Dinh  
; APPLICANT: Stuart, Susan G.  
; TITLE OF INVENTION: NOVEL SERPIN DERIVED FROM HUMAN  
; TITLE OF INVENTION: HYPOTHALAMUS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:



ADDRESS: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: US  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: PastSQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/203,237  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/997,040  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Luther, Barbara J.  
 REGISTRATION NUMBER: 33,954  
 REFERENCE/DOCKET NUMBER: PF-0039 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-855-0555  
 TELEFAX: 415-852-0195  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 382 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 US-09-203-237-3

Query Match 3.4%; Score 13; DB 2; Length 382;  
 Best Local Similarity 100.0%; Pred. No. 0.0004;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNAVYFKGW 181  
 Db 163 MVLNAVYFKGW 175

RESULT 13  
 US-07-768-286B-4  
 Sequence 4, Application US/07768286B  
 Patent No. 5444153  
 GENERAL INFORMATION:  
 APPLICANT: Goss, Neil H.  
 APPLICANT: RICHARDSON, Michael A.  
 TITLE OF INVENTION: VARIANTS OF PAL-2  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 1800 Diagonal Road, Suite 500  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: USA  
 ZIP: 22313-0299  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/768,2863  
 FILING DATE: 1991101  
 CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/AU90/00603  
 FILING DATE: 20-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 16786/157 CHAC  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 392 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-07-768-286B-4

Query Match 3.4%; Score 13; DB 1; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 0.00025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNAVYFKGW 181  
 Db 173 MVLNAVYFKGW 185

RESULT 14  
 US-08-121-714-7  
 Sequence 7, Application US/08121714  
 Patent No. 5470970  
 GENERAL INFORMATION:  
 APPLICANT: Sager, Ruth  
 TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH  
 TUMOR SUPPRESSING ACTIVITY  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson  
 STREET: 225 Franklin Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: U.S.A.  
 ZIP: 02110-2804  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTER: IBM PS/2 Model 502 or 55SX  
 OPERATING SYSTEM: MS-DOS (Version 5.0)  
 SOFTWARE: WordPerfect (Version 5.1)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/121,714  
 FILING DATE:  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/938,823  
 FILING DATE: 09/01/92  
 APPLICATION NUMBER: 07/844,296  
 FILING DATE: 02/28/92  
 APPLICATION NUMBER: 07/662,216  
 FILING DATE: 02/28/91  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fraser, Janis K.  
 REGISTRATION NUMBER: 34,819  
 REFERENCE/DOCKET NUMBER: 00530/072001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-5070  
 TELEFAX: (617) 542-8906  
 TELEX: 200154  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 405  
 TYPE: amino acid  
 STRANDEDNESS:

Search completed: December 12, 2003, 16:27:58  
Job time : 22 secs

TOPOLOGY: linear  
US-08-121-714-7  
Query Match 3.4%; Score 13; DB 1; Length 405;  
Best Local Similarity 100.0%; Pred.No. 0.00025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181  
Db 196 MVLNNAVYFKGW 208

RESULT 15  
US-08-477-108A-7  
; Sequence 7, Application US/08477108A  
; Patent No. 5801001  
; GENERAL INFORMATION:  
; APPLICANT: Sager, Ruth  
; APPLICANT: Zou, Zhigang  
; APPLICANT: Anisowicz, Anthony  
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH  
; TITLE OF INVENTION: TUNOR SUPPRESSING ACTIVITY  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IEM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,108A  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/121,714  
; FILING DATE: 09/01/93  
; APPLICATION NUMBER: 07/938,823  
; FILING DATE: 09/01/92  
; APPLICATION NUMBER: 07/844,296  
; FILING DATE: 02/28/92  
; APPLICATION NUMBER: 07/662,216  
; FILING DATE: 02/28/91  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 06570/002002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 405  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear

US-08-477-108A-7  
Query Match 3.4%; Score 13; DB 1; Length 405;  
Best Local Similarity 100.0%; Pred.No. 0.00025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181  
Db 196 MVLNNAVYFKGW 208

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:27:28 : Search time 43 Seconds  
(without alignments)  
1402.700 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 380

Sequence: 1 MASLAANAEEFCNLFREMD.....FLFVIRKDDIILFSGKVSQCP 380

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: A\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	D3	ID	Description
1	380	100.0	380	15	AA848379	Human megakaryocyte
2	380	100.0	380	20	AA848379	Human megaslin prote
3	380	100.0	380	21	AA848379	Human megaslin prote
4	380	100.0	380	22	AA848379	Human megaslin prote
5	380	100.0	380	24	AA848379	Lung cancer-associ
6	18	4.7	368	20	AA848379	Mouse megaslin prote
7	18	4.7	368	21	AA848379	Mouse megaslin prote
8	18	4.7	368	22	AA848379	Murine megaslin prote
9	18	4.7	368	20	AA848379	Rat megaslin protein

10 18 4.7 380 21 AA848379 Rat megaslin protein  
11 18 4.7 380 22 AA848379 Rat megaslin protein  
12 18 4.7 380 25 AA848379 Human megakaryocyte  
13 17 4.5 20 15 AA848379 Human megaslin prote  
14 16 4.2 16 21 AA848379 Megsln peptide 4  
15 16 4.2 16 21 AA848379 Human megaslin domai  
16 16 4.2 16 21 AA848379 Human megaslin domai  
17 15 3.9 15 20 AA848379 Megsln peptide 1  
18 15 3.9 15 20 AA848379 Megsln peptide 2  
19 15 3.9 15 20 AA848379 Megsln peptide 3  
20 15 3.9 15 21 AA848379 Human megaslin domai  
21 15 3.9 15 21 AA848379 Human megaslin domai  
22 15 3.9 15 22 AA848379 Human megaslin domai  
23 15 3.9 15 22 AA848379 Human megaslin domai  
24 13 3.4 361 9 AA848379 Sequence of plasmid  
25 13 3.4 382 12 AA848379 Plasmidogen activa  
26 13 3.4 392 12 AA848379 Plasmidogen activa  
27 13 3.4 415 8 AA848379 Sequence of human  
28 13 3.4 415 9 AA848379 Sequence of human  
29 13 3.4 415 10 AA848379 Encodes Plasmidogen  
30 13 3.4 415 12 AA848379 Placental PAI-2  
31 13 3.4 415 12 AA848379 PAI-2 from pDBP3.  
32 13 3.4 415 13 AA848379 Human plasminogen  
33 13 3.4 415 22 AA848379 Plasminogen activa  
34 13 3.4 415 23 AA848379 Human PAI-2 polyo  
35 13 3.4 423 22 AA848379 Human colon cancer  
36 13 3.4 434 12 AA848379 PreA-plasminogen a  
37 13 3.4 438 12 AA848379 PreB-PAI-2 with al  
38 13 3.4 438 12 AA848379 Human megakaryocyte  
39 11 2.9 17 15 AA848379 Human LSI-01 prote  
40 10 2.6 16 23 AA848379 Human megakaryocyte  
41 10 2.6 19 15 AA848379 Human megakaryocyte  
42 10 2.6 378 19 AA848379 Human protease nex  
43 10 2.6 378 19 AA848379 Human protease nex  
44 10 2.6 378 19 AA848379 Human protease nex  
45 10 2.6 378 19 AA848379 Human protease nex

#### ALIGNMENTS

RESULT 1  
AA848379  
ID AA848379 standard; Protein; 380 AA.  
XX  
AC AA848379;  
XX  
DT 25-MAR-2003 (updated)  
DT 16-AUG-1994 (first entry)  
XX  
DE Human megakaryocyte differentiation factor.  
XX  
KW Human megakaryocyte differentiation factor; MDF; thrombopoietin;  
KW haematopoietic stimulating factor; thrombocytopaenia; platelet;  
KW bone marrow transplantation; cancer chemotherapy.  
OS Homo sapiens.  
PN EP583884-AA1.  
XX  
PD 23-FEB-1994.  
XX  
PF 19-JUL-1993; 93EP-0305654.  
XX  
PR 17-JUL-1992; 92JP-0212305.  
PR 04-MAR-1993; 93JP-0067339.  
XX  
PA (SUNR) SUNTORY LTD.  
PA (TSUJ) TSUJIMOTO M.  
XX  
PI Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;  
PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamaichi K;  
XX

DR WPI; 1994-058782/08.  
 DR N-PSDB; AAQ56670.  
 PT New megakaryocyte differentiation factor - isolated from human  
 PT epidermoid carcinoma cells, used to treat conditions involving a  
 PT decrease in platelets  
 XX  
 PS Claim 7; Page 30-32; 47pp; English.  
 XX  
 CC Human MDF can be isolated from a culture of human epidermoid  
 CC carcinoma A431 cells in protein-free medium. The MDF stimulates  
 CC differentiation of megakaryocytes from myeloid cells in the presence  
 CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful  
 CC for treatment of diseases involving a decrease in platelet number  
 CC (esp. thrombocytopaenia) such as occurs in bone marrow  
 CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kD by  
 CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 380 AA;  
 Query Match 100.0%; Score 380; DB 15; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASLAANAARFCNLFREMDNQGNGNVFFSSLSFAALALVRLGAQDDSLQIDKLLHV 60  
 DB 1 MASLAANAARFCNLFREMDNQGNGNVFFSSLSFAALALVRLGAQDDSLQIDKLLHV 60  
 QY 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSTVNGLFAEKVYGFHKDYIECAEK 120  
 DB 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSTVNGLFAEKVYGFHKDYIECAEK 120  
 QY 121 LVDKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGEGG1SSSAVWLVNAVYFKGK 180  
 DB 121 LVDKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGEGG1SSSAVWLVNAVYFKGK 180  
 QY 181 WSAFTKSETINCHFKSPKCGKAVAMMHOERKPNLSVIEDPSMKILELYNGGINMVL 240  
 DB 181 WSAFTKSETINCHFKSPKCGKAVAMMHOERKPNLSVIEDPSMKILELYNGGINMVL 240  
 QY 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLALGLKIDIF 300  
 DB 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLALGLKIDIF 300  
 QY 301 DESKADLSGIASGGRLYISRMMHKSYLEVTEEGTATAATGSNIVEKQLPQSTLFRADHP 360  
 DB 301 DESKADLSGIASGGRLYISRMMHKSYLEVTEEGTATAATGSNIVEKQLPQSTLFRADHP 360  
 QY 361 FLFVIRKDDIILFSGKVSCP 380  
 DB 361 FLFVIRKDDIILFSGKVSCP 380  
 RESULT 2  
 ID AAY08254  
 AC AAY08254; standard; Protein; 380 AA.  
 XX  
 DT 14-JUL-1999 (first entry)  
 XX  
 DE Human megasin protein.  
 XX  
 KW Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;  
 KW human; rat; murine.  
 XX  
 OS Homo sapiens.  
 XX  
 FN W09915652-Al.  
 XX  
 PD 01-APR-1999.  
 XX

PF 22-SEP-1998; 98WO-JPC4269.  
 XX  
 PR 22-SEP-1997; 97JP-0275302.  
 XX  
 PA (KURO/) KUROKAWA K.  
 PA (MIYA/) MIYATA T.  
 XX  
 PI Miyata T;  
 XX  
 DR WPI; 1999-276983/23.  
 DR N-PSDB; AA56712.  
 XX  
 PT Megsin protein expressed specifically in mesangial cells  
 PS Claim 1; Page 62-64; 100pp; Japanese.  
 CC This invention describes the isolation of novel megsin nucleic acid and  
 CC proteins from human, rat and mouse tissue. This protein is expressed  
 CC specifically in mesangial cells. The products of the invention are  
 CC useful for the treatment and diagnosis of diseases involving mesangial  
 CC cells, such as IgA nephropathy.  
 XX  
 SQ Sequence 380 AA;  
 Query Match 100.0%; Score 380; DB 20; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASLAANAARFCNLFREMDNQGNGNVFFSSLSFAALALVRLGAQDDSLQIDKLLHV 60  
 DB 1 MASLAANAARFCNLFREMDNQGNGNVFFSSLSFAALALVRLGAQDDSLQIDKLLHV 60  
 QY 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSTVNGLFAEKVYGFHKDYIECAEK 120  
 DB 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSTVNGLFAEKVYGFHKDYIECAEK 120  
 QY 121 LVDKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGEGG1SSSAVWLVNAVYFKGK 180  
 DB 121 LVDKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGEGG1SSSAVWLVNAVYFKGK 180  
 QY 181 WSAFTKSETINCHFKSPKCGKAVAMMHOERKPNLSVIEDPSMKILELYNGGINMVL 240  
 DB 181 WSAFTKSETINCHFKSPKCGKAVAMMHOERKPNLSVIEDPSMKILELYNGGINMVL 240  
 QY 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLALGLKIDIF 300  
 DB 241 LPENDLSIENKLTFOQLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLALGLKIDIF 300  
 QY 301 DESKADLSGIASGGRLYISRMMHKSYLEVTEEGTATAATGSNIVEKQLPQSTLFRADHP 360  
 DB 301 DESKADLSGIASGGRLYISRMMHKSYLEVTEEGTATAATGSNIVEKQLPQSTLFRADHP 360  
 QY 361 FLFVIRKDDIILFSGKVSCP 380  
 DB 361 FLFVIRKDDIILFSGKVSCP 380  
 RESULT 3  
 ID AAB24142  
 AC AAB24142; standard; Protein; 380 AA.  
 XX  
 DT 30-JAN-2001 (first entry)  
 XX  
 DE Human megsin protein sequence SEQ ID NO:2.  
 KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;  
 KW IgA; immunoglobulin A; detection; renal function; renal disorder;  
 KW diagnosis; biological sample; blood; urine.  
 OS Homo sapiens.  
 XX

FN W0200057189-A1.

XX 28-SEP-2000.

XX 17-MAR-2000; 2000WO-JP01646.

XX 19-MAR-1999; 99JP-0075305.

XX 28-OCT-1999; 99JP-0306623.

XX (KURO/) KUROKAWA K.

XX (FUSO ) FUSO PHARM IND LTD.

XX (MIYA/) MIYATA T.

XX Miyata T;

XX WPI; 2000-611642/58.

XX N-PSDB; AAA99294.

XX Evaluating renal function comprises assaying megin protein in

XX biological sample -

XX Example 2; Page 66-69; 93pp; Japanese.

XX The present invention describes a method for evaluating renal function.

XX The method comprises assaying megin protein in biological sample. Also

XX described are: (1) use of an anti-megsin protein antibody for diagnosing

XX renal function; and (2) a kit for detecting megin protein comprising:

XX (a) anti-megsin protein antibody attached to solid magnetic particles;

XX (b) direct or indirect fixing for the antibody to the particles; and

XX (c) a magnet. The process is useful for evaluating renal function and

XX diagnosing renal disorders by assaying megin protein in biological

XX samples (preferably urine or blood). The process is reproducible and

XX gives accurate results. The present sequence represents the human megin

XX protein, which is given in the exemplification of the present invention.

XX Sequence 380 AA;

XX Query Match 100.0%; Score 380; DB 21; Length 380;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAECFNLPREMDNQGNGNVFFSSLSLFAALALVRLGAQDDLSQIDKLLHV 60

Db 1 MASLAANAECFNLPREMDNQGNGNVFFSSLSLFAALALVRLGAQDDLSQIDKLLHV 60

QY 61 NTASGYGNSNSQSGLSQOLKEVPSDINASHKDYDLSIVNGLFAEKVGFPHDYIECAEK 120

Db 61 NTASGYGNSNSQSGLSQOLKEVPSDINASHKDYDLSIVNGLFAEKVGFPHDYIECAEK 120

QY 121 LYDAKVERVDFTNHLEDTRNINKWENETHGKIKNVI GEGGSSSAVMVLYNAVYFKGK 180

Db 121 LYDAKVERVDFTNHLEDTRNINKWENETHGKIKNVI GEGGSSSAVMVLYNAVYFKGK 180

QY 181 WSAFTKSETINCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKILELRYNGGNNMVL 240

Db 181 WSAFTKSETINCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKILELRYNGGNNMVL 240

QY 241 LPENDLSEIENKLTFTQNLMEWNTNPRMTSKYVEVFFPOFKIEKNYEMKYLALGLKIDIF 300

Db 241 LPENDLSEIENKLTFTQNLMEWNTNPRMTSKYVEVFFPOFKIEKNYEMKYLALGLKIDIF 300

QY 301 DESKADLSGIASGGRLYISRMMHKSYLEVTEEGTEATATGTSNIVEKQLPQSTLFRADHP 360

Db 301 DESKADLSGIASGGRLYISRMMHKSYLEVTEEGTEATATGTSNIVEKQLPQSTLFRADHP 360

QY 361 FLFVIRKDDIILFSGKVSCP 380

Db 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 4

AAB83075

TD AAB83075 standard; Protein; 380 AA.

XX

AC AAB83075;

XX 10-JUL-2001 (first entry)

XX Human megin protein.

XX Human; megin; mesangial cell proliferative nephritis; nephrotropic;

XX transgenic mouse; glomerular disease; animal model; drug screening.

XX Homo sapiens.

XX W0200124628-A1.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-JP06988.

XX 06-OCT-1999; 99JP-0285736.

XX (KURO/) KUROKAWA K.

XX (MIYA/) MIYATA T.

XX Miyata T;

XX WPI; 2001-300136/31.

XX N-PSDB; AAF82438.

XX Mouse model for mesangial cell proliferative nephritis for development

XX and screening of new treatments -

XX Example 4; Page 44-46; 62pp; Japanese.

XX The present sequence is human megin. The human megin coding

XX sequence may be introduced into a mouse to produce an animal model of

XX mesangial cell proliferative nephritis. The symptoms include

XX enlargement of the mesangial base region, sedimentation of an immune

XX complex and an increase in mesangial cells. The animal model is useful

XX for analysing the pathology of chronic glomerular diseases and for

XX screening compositions for prevention and treatment of the diseases.

XX Highly uniform models can be made easily and in large numbers using

XX this method.

XX Sequence 380 AA;

XX Query Match 100.0%; Score 380; DB 22; Length 380;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAECFNLPREMDNQGNGNVFFSSLSLFAALALVRLGAQDDLSQIDKLLHV 60

Db 1 MASLAANAECFNLPREMDNQGNGNVFFSSLSLFAALALVRLGAQDDLSQIDKLLHV 60

QY 61 NTASGYGNSNSQSGLSQOLKEVPSDINASHKDYDLSIVNGLFAEKVGFPHDYIECAEK 120

Db 61 NTASGYGNSNSQSGLSQOLKEVPSDINASHKDYDLSIVNGLFAEKVGFPHDYIECAEK 120

QY 121 LYDAKVERVDFTNHLEDTRNINKWENETHGKIKNVI GEGGSSSAVMVLYNAVYFKGK 180

Db 121 LYDAKVERVDFTNHLEDTRNINKWENETHGKIKNVI GEGGSSSAVMVLYNAVYFKGK 180

QY 181 WSAFTKSETINCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKILELRYNGGNNMVL 240

Db 181 WSAFTKSETINCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKILELRYNGGNNMVL 240

QY 241 LPENDLSEIENKLTFTQNLMEWNTNPRMTSKYVEVFFPOFKIEKNYEMKYLALGLKIDIF 300

Db 241 LPENDLSEIENKLTFTQNLMEWNTNPRMTSKYVEVFFPOFKIEKNYEMKYLALGLKIDIF 300

QY 301 DESKADLSGIASGGRLYISRMMHKSYLEVTEEGTEATATGTSNIVEKQLPQSTLFRADHP 360

Db 301 DESKADLSGIASGGRLYISRMMHKSYLEVTEEGTEATATGTSNIVEKQLPQSTLFRADHP 360

QY 361 FLFVIRKDDIIILFSGKVSCP 380  
 DB 361 FLFVIRKDDIIILFSGKVSCP 380

RESULT 5  
 ABUS6533  
 ID ABUS6533 standard; Protein; 380 AA.  
 XX  
 AC ABUS6533;  
 XX  
 DT 02-APR-2003 (first entry)  
 XX  
 DE Lung cancer-associated polypeptide #126.  
 XX  
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
 KW antinflammatory; anasthmatic; non-small cell lung cancer; atelectasis;  
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
 XX  
 OS Unidentified.  
 XX  
 FN WO200286443-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 18-APR-2002; 2002WO-US12476.  
 XX  
 PR 18-APR-2001; 2001US-284770P.  
 PR 10-MAY-2001; 2001US-290492P.  
 PR 09-NOV-2001; 2001US-339245P.  
 PR 13-NOV-2001; 2001US-350666P.  
 PR 29-NOV-2001; 2001US-334370P.  
 PR 12-APR-2002; 2002US-372246P.  
 XX  
 PA (EOSE-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Aziz N, Murray R;  
 XX  
 DR WFI; 2003-093161/08.  
 DR N-PSDB; ABX76260.

Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer -

Claim 27; Page 288; 453pp; English.

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the invention.

Sequence 380 AA;

Query Match

100.0%; Score 380; DB 24; Length 380;

Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAARFCNLFREMDNQGNGVFFSSLSLFAALALVRIGAGQDLSQIDKLLHV 60  
 DB 1 MASLAANAARFCNLFREMDNQGNGVFFSSLSLFAALALVRIGAGQDLSQIDKLLHV 60

QY 61 NTASGYGNSSNSQSLQSLXRVFSDINASHKDYDLSTVNGLFPAEKYGFHDKYTECAEK 120  
 DB 61 NTASGYGNSSNSQSLQSLXRVFSDINASHKDYDLSTVNGLFPAEKYGFHDKYTECAEK 120

QY 121 LYDAKVERVDFTNHLEDTNRNINKWVENETHGKIKNVIGEGGSSSAVAVLVNAVYFKGK 180  
 DB 121 LYDAKVERVDFTNHLEDTNRNINKWVENETHGKIKNVIGEGGSSSAVAVLVNAVYFKGK 180

QY 181 QWSAFTKSETINCHFKPSKSGKAVAMHQRKFNLSVIEDPSMKLELRNGGINMYVL 240  
 DB 181 QWSAFTKSETINCHFKPSKSGKAVAMHQRKFNLSVIEDPSMKLELRNGGINMYVL 240

QY 241 LPENDLSEIENKLTQNLMEWTNPRMTSKYVEVFFPQFKIEKNYEMKQYLFALGLKDI 300  
 DB 241 LPENDLSEIENKLTQNLMEWTNPRMTSKYVEVFFPQFKIEKNYEMKQYLFALGLKDI 300

QY 301 DESKADLSGIASGRLYISRMWHKSYTEVTERTGTEATAATGSNIVEKQLPOSTILFRADHP 360  
 DB 301 DESKADLSGIASGRLYISRMWHKSYTEVTERTGTEATAATGSNIVEKQLPOSTILFRADHP 360

QY 361 FLFVIRKDDIIILFSGKVSCP 380  
 DB 361 FLFVIRKDDIIILFSGKVSCP 380

## RESULT 6

AAU08256  
 ID AAU08256 standard; Protein; 368 AA.

XX AC AAU08256;  
 XX DT 14-JUL-1999 (first entry)  
 XX DE Mouse megasin protein.

XX KW Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;  
 XX KW human; rat; murine.

XX OS Mus musculus.

XX FH Key  
 XX FT Protein  
 XX FT Location/Qualifiers  
 1..368  
 /note= "partial sequence"

XX FN WO9915652-A1.

XX PD 01-APR-1999.

XX PF 22-SEP-1998; 98WO-JP04269.

XX PR 22-SEP-1997; 97JP-0275302.

XX PA (KURO/) KUROKAWA K.

XX PA (MIYA/) MIYATA T.

XX PI Miyata T;

XX DR WFI; 1999-275983/23.

XX DR N-PSDB; AAU56714.

XX PT Megsin protein expressed specifically in mesangial cells

XX PS Claim 1; Page 76-79; 100pp; Japanese.

XX CC This invention describes the isolation of novel megasin nucleic acid and  
 CC proteins from human, rat and mouse tissue. This protein is expressed

CC specifically in mesangial cells. The products of the invention are  
 CC useful for the treatment and diagnosis of diseases involving mesangial  
 CC cells, such as IGA nephropathy.  
 XX  
 SQ

Sequence 368 AA;

Query Match 4.7%; Score 18; DB 20; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVMLVNAVYFKGW 181  
 DB 152 SSSAVMLVNAVYFKGW 169

# RESULT 7

AAE24151  
 ID AAE24151 standard; Protein; 368 AA.

XX AC AAE24151;

XX DT 30-JAN-2001 (first entry)

DE Mouse megin protein sequence SEQ ID NO:21.

XX Megin; mesangium-predominant gene; serpin regulated; nephropathy;  
 KW Iga; immunoglobulin A; detection; renal function; renal disorder;  
 KW diagnosis; biological sample; blood; urine.  
 XX OS Mus musculus.

XX PN WO200057189-A1.

XX PD 28-SEP-2000.

XX PF 17-MAR-2000; 2000WO-JP01646.

XX PR 19-MAR-1999; 99JP-0075305.

XX PR 28-OCT-1999; 99JP-0306623.

XX PA (KURO/) KUROKAWA K.

XX PA (FUSO ) FUSO PHARM IND LTD.

XX PA (MIYA/) MIYATA T.

XX PI Miyata T;

XX DR WPI; 2000-611642/56.

XX DR N-PSDB; AAC55239.

XX PT Evaluating renal function comprises assaying megin protein in  
 PT biological sample -

PS Disclosure; Page 89-91; 93pp; Japanese.

CC The present invention describes a method for evaluating renal function.  
 CC The method comprises assaying megin protein in biological sample. Also  
 CC described are: (1) use of an anti-megin protein antibody for diagnosing  
 CC renal function; and (2) a kit for detecting megin protein comprising:  
 CC (a) anti-megin protein antibody attached to solid magnetic particles;  
 CC (b) direct or indirect fixing for the antibody to the particles; and  
 CC (c) a magnet. The process is useful for evaluating renal function and  
 CC diagnosing renal disorders by assaying megin protein in biological  
 CC samples (preferably urine or blood). The process is reproducible and  
 CC gives accurate results. The present sequence represents the mouse megin  
 CC protein, which is given in the exemplification of the present invention.  
 XX Sequence 368 AA;

Query Match 4.7%; Score 18; DB 21; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVMLVNAVYFKGW 181

DB 152 SSSAVMLVNAVYFKGW 169

# RESULT 8

AAE83077

ID AAE83077 standard; Protein; 368 AA.

XX AC AAE83077;

XX DT 10-JUL-2001 (first entry)

DE Murine megin protein.

XX Mouse; megin; mesangial cell proliferative nephritis; nephrotropic;  
 KW transgenic mouse; glomerular disease; animal model; drug screening.  
 XX OS Mus musculus.

XX PN WO200124628-A1.

XX PD 12-APR-2001.

XX PF 06-OCT-2000; 2000WO-JP06988.

XX PR 06-OCT-1999; 99JP-0285736.

XX PA (KURO/) KUROKAWA K.

XX PA (MIYA/) MIYATA T.

XX PI Miyata T;

XX DR WPI; 2001-300136/31.

XX DR N-PSDB; AAF82440.

XX PT Mouse model for mesangial cell proliferative nephritis for development  
 PT and screening of new treatments -

PS Disclosure; Page 52-53; 62pp; Japanese.

XX The present sequence is murine megin protein. The human megin coding  
 CC sequence may be introduced into a mouse to produce an animal model of  
 CC mesangial cell proliferative nephritis. The symptoms include  
 CC enlargement of the mesangial base region, sedimentation of an immune  
 CC complex and an increase in mesangial cells. The animal model is useful  
 CC for analysing the pathology of chronic glomerular diseases and for  
 CC screening compositions for prevention and treatment of the diseases.  
 CC Highly uniform models can be made easily and in large numbers using  
 CC this method.

XX Sequence 368 AA;

Query Match 4.7%; Score 18; DB 22; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVMLVNAVYFKGW 181

DB 152 SSSAVMLVNAVYFKGW 169

# RESULT 9

AAV08255

ID AAV08255 standard; Protein; 380 AA.

XX AC AAV08255;

XX DT 14-JUL-1999 (first entry)

XX DE Rat megin protein.

XX Megin; mesangial cell; treatment; diagnosis; disease; IGA nephropathy;  
 KW human; rat; murine.

```

XX OS Rattus rattus.
XX PI WO9915652-A1.
XX PN 01-APR-1999.
XX PD 22-SEP-1998; 98WO-JP04269.
XX PE 22-SEP-1997; 97JP-0275302.
XX PF (KURO/) KUROKAWA K.
XX PG (MIYA/) MIYATA T.
XX PH Miyata T.
XX PI WPI; 1999-276983/23.
XX PJ N-PSDB; AAX56712.
XX PK Megsin protein expressed specifically in mesangial cells
XX PL Claim 1; Page 69-72; 100pp; Japanese.
XX PM This invention describes the isolation of novel megin nucleic acid and
XX PN proteins from human, rat and mouse tissue. This protein is expressed
XX PO specifically in mesangial cells. The products of the invention are
XX PP useful for the treatment and diagnosis of diseases involving mesangial
XX PQ cells, such as IgA nephropathy.
XX PR
XX PS Sequence 380 AA;
XX SQ
Query Match 4.7%; Score 18; DB 20; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVVVLVNAVYFKGKW 181
DB 164 SSSAVVVLVNAVYFKGKW 181

RESULT 10
AAB24150
ID AAB24150 standard; Protein; 380 AA.
XX AC AAB24150;
XX DT 30-JAN-2001 (first entry)
XX DE Rat megin protein sequence SEQ ID NO:19.
XX KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
XX KW IgA; immunoglobulin A; detection; renal function; renal disorder;
XX KW diagnosis; biological sample; blood; urine.
XX OS Rattus norvegicus.
XX PH Key Location/Qualifiers
XX PI Misc-difference 51
XX PT Misc-difference 94
XX FT Misc-difference 94
XX FT /note= "unspecified"
XX FT /note= "unspecified"
XX PN WO200057189-A1.
XX PD 28-SEP-2000.
XX PF 17-MAR-2000; 2000WO-JP01646.
XX PG 19-MAR-1999; 99JP-0075305.
XX PR 28-OCT-1999; 99JP-0306623.
XX PY (KURO/) KUROKAWA K.
XX PA (FUSO ) FUSO PHARM IND LTD.

PA (MIYA/) MIYATA T.
XX PI Miyata T.
XX PN WPI; 2000-611642/58.
XX PD N-PSDB; AAC55238.
XX PE Evaluating renal function comprises assaying megin protein in
XX PF biological sample -
XX PG Example 2; Page 81-84; 93pp; Japanese.
XX PH The present invention describes a method for evaluating renal function.
XX PI The method comprises assaying megin protein in biological sample. Also
XX PJ described are: (1) use of an antimegin protein antibody for diagnosing
XX PK renal function; and (2) a kit for detecting megin protein comprising:
XX PL (a) antimegin protein antibody attached to solid magnetic particles;
XX PM (b) direct or indirect fixing for the antibody to the particles; and
XX PN (c) a magnet. The process is useful for evaluating renal function and
XX PO diagnosing renal disorders by assaying megin protein in biological
XX PP samples (preferably urine or blood). The process is reproducible and
XX PQ gives accurate results. The present sequence represents the rat megin
XX PR protein, which is given in the exemplification of the present invention.
XX PS Sequence 380 AA;
XX SQ
Query Match 4.7%; Score 18; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVVVLVNAVYFKGKW 181
DB 164 SSSAVVVLVNAVYFKGKW 181

RESULT 11
AAG64286
ID AAG64286 standard; Protein; 380 AA.
XX AC AAG64286;
XX DT 21-SEP-2001 (first entry)
XX DE Rat megin protein.
XX KW Rat; megin; renal mesangial cell; mesangium proliferative nephritis.
XX OS Rattus norvegicus.
XX PN WO200148019-A1.
XX PD 05-JUL-2001.
XX PF 26-DEC-2000; 2000WO-JP09251.
XX PG 28-DEC-1999; 99JP-0373677.
XX PH (KURO/) KUROKAWA K.
XX PI (MIYA/) MIYATA T.
XX PN Miyata T.
XX PD WPI; 2001-425651/45.
XX PR N-PSDB; AAX48181.
XX PT New antibody recognizing a partial sequence of rat megin protein for
XX PG diagnosis of mesangium proliferative nephritis -
XX PS Disclosure; Page 54-56; 63pp; Japanese.
XX CC The present invention relates to a novel antibody which recognises a
XX CC peptide consisting of residues 341-354 of rat megin protein. The present
XX CC sequence is the protein sequence for rat megin, which was used in the

```



CC present invention. Mesgin is highly expressed in renal mesangial cells  
 CC and its level is elevated in mesangium proliferative nephritis. Assay of  
 CC the serum or urine level using the antibody is therefore indicative of  
 CC this type of disorder.

XX Sequence 380 AA;

Query Match 4.7%; Score 18; DB 22; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWLVNVAVFKGW 181  
 |||||  
 DB 164 SSSAVWLVNVAVFKGW 181

## RESULT 12

AA83076  
 ID AAB83076 standard; Protein; 380 AA.

XX  
 AC AAB83076;

DT 10-JUL-2001 (first entry)

DE Rat mesgin protein.

XX Rat; mesgin; mesangial cell proliferative nephritis; nephrotropic;  
 KW transgenic mouse; glomerular disease; animal model; drug screening.  
 XX Rattus norvegicus.

PN WO200124628-A1.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-JP06988.

XX 06-OCT-1999; 95JP-0285736.

XX (KURO/) KUROKAWA K.  
 PA (MIYA/) MIYATA T.

XX Miyata T;

DR WPI: 2001-300136/31.

DR N-PSDB; AAB82439.

XX Mouse model for mesangial cell proliferative nephritis for development  
 PT and screening of new treatments -

XX Disclosure; Page 48-50; 62pp; Japanese.

XX The present sequence is rat mesgin. The human mesgin coding  
 CC sequence may be introduced into a mouse to produce an animal model of  
 CC mesangial cell proliferative nephritis. The symptoms include  
 CC enlargement of the mesangial base region, sedimentation of an immune  
 CC complex and an increase in mesangial cells. The animal model is useful  
 CC for analysing the pathology of chronic glomerular diseases and for  
 CC screening compositions for prevention and treatment of the diseases.  
 CC Highly uniform models can be made easily and in large numbers using  
 CC this method.

XX Sequence 380 AA;

Query Match 4.7%; Score 18; DB 22; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWLVNVAVFKGW 181  
 |||||  
 DB 164 SSSAVWLVNVAVFKGW 181

## RESULT 13

AA857112  
 ID AAR57112 standard; peptide; 20 AA.

XX  
 AC AAR57112;

DT 25-MAR-2003 (updated)

DT 16-AUG-1994 (first entry)

DE Human megakaryocyte differentiation factor peptide 6.

XX Human megakaryocyte differentiation factor; MDF; thrombopoietin;  
 KW haematopoietic stimulating factor; thrombocytopaenia; platelet;  
 KW bone marrow transplantation; cancer chemotherapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 18

FT /note= "not determined"

XX EP583884-A1.

XX 23-FEB-1994.

XX 19-JUL-1993; 93EP-0305654.

XX 17-JUL-1992; 92JP-0212305.

XX 04-MAR-1993; 93JP-0067339.

XX (SUNR) SUNTORY LTD.

XX (TSUJ/) TSUJIMOTO M.

XX Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;

PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamaichi K;

XX WPI: 1994-058782/08.

XX New megakaryocyte differentiation factor - isolated from human  
 PT epidermoid carcinoma cells, used to treat conditions involving a  
 PT decrease in platelets

XX Claim 1; Page 20; 47pp; English.

XX Human MDF (see AAR48379) can be isolated from a culture of human  
 CC epidermoid carcinoma A431 cells in protein-free medium. The MDF  
 CC stimulates differentiation of megakaryocytes from myeloid cells  
 CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin  
 CC making it useful for treatment of diseases involving a decrease  
 CC in platelet number (esp. thrombocytopaenia) such as occurs in bone  
 CC marrow transplantation and in chemotherapy. MDF has mol.wt.  
 CC 55-57kD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and  
 CC contains an amino acid sequence comprising at least one of the  
 CC sequences AAR57107-R57115.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 20 AA;

Query Match 4.5%; Score 17; DB 15; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIAGSGRLYISRM 321

DB 1 ADLSGIAGSGRLYISRM 17

## RESULT 14

AA08260  
 ID AAY08260 standard; Protein; 16 AA.

XX  
 AC AAY08260;

DT 14-JUL-1999 (first entry)  
 XX  
 DE Megin peptide 4.  
 XX  
 KW Megin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;  
 XX human; rat; murine.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9915652-A1.  
 XX  
 PD 01-APR-1999.  
 XX  
 PF 22-SEP-1998; 98WO-JF04269.  
 XX  
 PR 22-SEP-1997; 97JP-0275302.  
 XX  
 PA (KURO/) KUROKAWA K.  
 XX (MIYA/) MIYATA T.  
 XX  
 PI Miyata T;  
 XX  
 DR WPI; 1999-276983/23.  
 XX  
 PT Megin protein expressed specifically in mesangial cells  
 XX  
 PS Example 10; Page 86; 100pp; Japanese.  
 XX  
 CC This invention describes the isolation of novel megin nucleic acid and  
 CC proteins from human, rat and mouse tissue. This protein is expressed  
 CC specifically in mesangial cells. The products of the invention are  
 CC useful for the treatment and diagnosis of diseases involving mesangial  
 CC cells, such as IgA nephropathy. AA08257-Y08260 represent peptides used  
 CC in the method of the invention.  
 XX  
 SQ Sequence 16 AA;  
 Query Match 4.2%; Score 16; DB 20; Length 16;  
 Best Local Similarity 100.0%; Pred.No.1.3e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 339 ATGSNIVEKQLPQSTL 354  
 |||||  
 DB 1 ATGSNIVEKQLPQSTL 16  
 |||||  
 RESULT 15  
 AAB24145  
 ID AAB24145 standard; Peptide; 16 AA.  
 XX  
 AC AAB24145;  
 XX  
 DT 30-JAN-2001 (first entry)  
 XX  
 DE Human megin domain peptide SEQ ID NO:13.  
 XX  
 KW Megin; mesangium-predominant gene; serpin regulated; nephropathy;  
 XX IgA; immunoglobulin A; detection; renal function; renal disorder;  
 KW diagnosis; biological sample; blood; urine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2000057189-A1.  
 XX  
 PD 28-SEP-2000.  
 XX  
 PF 17-MAR-2000; 2000WO-JP01646.  
 XX  
 PR 19-MAR-1999; 99JP-0075305.  
 XX  
 PR 28-OCT-1999; 99JP-0306623.  
 XX  
 PA (KURO/) KUROKAWA K.  
 XX (FUSO) FUSO PHARM IND LTD.

PA (MIYA/) MIYATA T.  
 XX  
 PI Miyata T;  
 XX  
 DR WPI; 2000-611642/58.  
 XX  
 PT Evaluating renal function comprises assaying megin protein in  
 PT biological sample -  
 XX  
 PS Example 2; Page 29; 93pp; Japanese.  
 XX  
 CC The present invention describes a method for evaluating renal function.  
 CC The method comprises assaying megin protein in biological sample. Also  
 CC described are: (1) use of an antimegin protein antibody for diagnosing  
 CC renal function; and (2) a kit for detecting megin protein comprising:  
 CC (a) antimegin protein antibody attached to solid magnetic particles;  
 CC (b) direct or indirect fixing for the antibody to the particles; and  
 CC (c) a magnet. The process is useful for evaluating renal function and  
 CC diagnosing renal disorders by assaying megin protein in biological  
 CC samples (preferably urine or blood). The process is reproducible and  
 CC gives accurate results. The present sequence represents a human megin  
 CC domain peptide, which is used in an example from the present invention.  
 XX  
 SQ Sequence 16 AA;  
 Query Match 4.2%; Score 16; DB 21; Length 16;  
 Best Local Similarity 100.0%; Pred.No.1.3e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 339 ATGSNIVEKQLPQSTL 354  
 |||||  
 DB 1 ATGSNIVEKQLPQSTL 16  
 |||||  
 Search completed: December 12, 2003, 16:33:18  
 Job time : 44 secs



;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 4-212305  
;; FILING DATE: 17-JUL-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 6-067339  
;; FILING DATE: 04-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McGowan, Malcolm K.  
;; REGISTRATION NUMBER: 39,300  
;; REFERENCE/DOCKET NUMBER: 001560-247  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 836-6620  
;; TELEFAX: (703) 836-2021  
;; INFORMATION FOR SEQ ID NO: 34:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 380 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-09-140-719-34

Query Match 100.0%; Score 380; DB 9; Length 380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASLAANAEEFCNLFREMDNQGNGNVFSSLSFAALALVRLGQDDSLSQIDKLLHV 60  
Db 1 MASLAANAEEFCNLFREMDNQGNGNVFSSLSFAALALVRLGQDDSLSQIDKLLHV 60  
Qy 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSVNGLFAEKVYGFHKDYIECAEK 120  
Db 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSVNGLFAEKVYGFHKDYIECAEK 120  
Qy 121 LYDAKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGEGGISSAVMVLVNAVYFKGK 180  
Db 121 LYDAKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGEGGISSAVMVLVNAVYFKGK 180  
Qy 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKPNLSVIEDPSMKILELYNGGINMVL 240  
Db 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKPNLSVIEDPSMKILELYNGGINMVL 240  
Qy 241 LPENDLSEIENKLTFFONLMEWTPRRMTSKYVEVFFPFQFKIEKNYEMKQYLFALGLKIDIF 300  
Db 241 LPENDLSEIENKLTFFONLMEWTPRRMTSKYVEVFFPFQFKIEKNYEMKQYLFALGLKIDIF 300  
Qy 301 DESKADLSGIASGRLYISRMMHKSYLEVTEEGTATAATGNSNIVEKQLPQSTLFRADHP 360  
Db 301 DESKADLSGIASGRLYISRMMHKSYLEVTEEGTATAATGNSNIVEKQLPQSTLFRADHP 360  
Qy 361 FLFVIRKDDIILFSGKVSCP 380  
Db 361 FLFVIRKDDIILFSGKVSCP 380

## RESULT 2

US-10-091-442-34

; Sequence 34, Application US/10091442

; Publication No. US20020164711A1

; GENERAL INFORMATION:

; APPLICANT: TSUJIMOTO, Masafumi

; TWASA, Fuyuki

; TSURUOKA, No. US20020164711A1

; NAKAZATO, Hiroshi

; MIURA, Kenju

; ISHIDA, No. US20020164711A1

; KURIHARA, Tatsuya

; YAMAUCHI, Kozo

; YAMAGUCHI, No. US20020164711A1

; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Burns, Doane, Swecker & Mathis

; STREET: P.O. Box 1404

;; CITY: Alexandria  
;; STATE: Virginia  
;; COUNTRY: United States  
;; ZIP: 22313-1404  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/091,442  
;; FILING DATE: 07-Mar-2002  
;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US/09/140,719

;; FILING DATE: 08-AUG-1998

;; APPLICATION NUMBER: US 08/474,661

;; FILING DATE: 07-JUN-1995

;; APPLICATION NUMBER: US 08/091,028

;; FILING DATE: 14-JUL-1993

;; APPLICATION NUMBER: JP 4-212305

;; FILING DATE: 17-JUL-1992

;; APPLICATION NUMBER: JP 6-067339

;; FILING DATE: 04-MAR-1993

;; ATTORNEY/AGENT INFORMATION:

;; NAME: McGowan, Malcolm K.

;; REGISTRATION NUMBER: 39,300

;; REFERENCE/DOCKET NUMBER: 001560-247

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (703) 836-6620

;; TELEFAX: (703) 836-2021

;; INFORMATION FOR SEQ ID NO: 34:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 380 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; SEQUENCE DESCRIPTION: SEQ ID NO: 34:

;; US-10-091-442-34

Query Match 100.0%; Score 380; DB 14; Length 380;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASLAANAEEFCNLFREMDNQGNGNVFSSLSFAALALVRLGQDDSLSQIDKLLHV 60

Db 1 MASLAANAEEFCNLFREMDNQGNGNVFSSLSFAALALVRLGQDDSLSQIDKLLHV 60

Qy 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSVNGLFAEKVYGFHKDYIECAEK 120

Db 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSVNGLFAEKVYGFHKDYIECAEK 120

Qy 121 LYDAKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGEGGISSAVMVLVNAVYFKGK 180

Db 121 LYDAKVERVDFTHLEDTNRNINKWVENETHGKIKNVIGEGGISSAVMVLVNAVYFKGK 180

Qy 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKPNLSVIEDPSMKILELYNGGINMVL 240

Db 181 WSAFTKSETINCHFKSPKCSKAVAMHQRKPNLSVIEDPSMKILELYNGGINMVL 240

Qy 241 LPENDLSEIENKLTFFONLMEWTPRRMTSKYVEVFFPFQFKIEKNYEMKQYLFALGLKIDIF 300

Db 241 LPENDLSEIENKLTFFONLMEWTPRRMTSKYVEVFFPFQFKIEKNYEMKQYLFALGLKIDIF 300

Qy 301 DESKADLSGIASGRLYISRMMHKSYLEVTEEGTATAATGNSNIVEKQLPQSTLFRADHP 360

Db 301 DESKADLSGIASGRLYISRMMHKSYLEVTEEGTATAATGNSNIVEKQLPQSTLFRADHP 360

Qy 361 FLFVIRKDDIILFSGKVSCP 380

Db 361 FLFVIRKDDIILFSGKVSCP 380

## RESULT 3

US-09-140-719-3  
; Sequence 3, Application US/09140719  
; Patent No. US20010026931A1  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Fuyuki  
; APPLICANT: TSUROOKA, No. US20010026931A1uo  
; APPLICANT: NAKAZATO, Hiroshi  
; APPLICANT: MIURA, Kenju  
; APPLICANT: ISHIDA, No. US20010026931A1uhiro  
; APPLICANT: KURIHARA, Tatsuya  
; APPLICANT: YAMAICHI, Kozo  
; APPLICANT: YAMAGUCHI, No. US20010026931A1omi  
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/140,719  
; FILING DATE: 08-AUG-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/474,661  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/091,028  
; FILING DATE: 14-JUL-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 4-212305  
; FILING DATE: 17-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-067339  
; FILING DATE: 04-WAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McGowan, Malcolm K.  
; REGISTRATION NUMBER: 39,300  
; REFERENCE/DOCKET NUMBER: 001560-247  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-140-719-3

Query Match 5.0%; Score 19; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.5e-11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 VERVDFTHLEEDTRRNK 144  
DB 1 VERVDFTHLEEDTRRNK 19

## RESULT 4

US-10-091-442-3  
; Sequence 3, Application US/10091442  
; Publication No. US20020164711A1

## GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSUROOKA, No. US20020164711A1uo  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. US20020164711A1uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAICHI, Kozo  
APPLICANT: YAMAGUCHI, No. US20020164711A1omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/091,442  
FILING DATE: 07-Mar-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/140,719  
FILING DATE: 08-AUG-1998  
APPLICATION NUMBER: US 08/474,661  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
APPLICATION NUMBER: JP 6-067339  
FILING DATE: 04-WAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-247  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-091-442-3

Query Match 5.0%; Score 19; DB 14; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.5e-11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 VERVDFTHLEEDTRRNK 144  
DB 1 VERVDFTHLEEDTRRNK 19

## RESULT 5

US-09-140-719-6  
; Sequence 6, Application US/09140719  
; Patent No. US20010026931A1  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Fuyuki  
; APPLICANT: TSUROOKA, No. US20010026931A1uo

APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. US20010026931A1uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAICHI, Kozo  
APPLICANT: YAMAGUCHI, No. US20010026931A1omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/140,719  
FILING DATE: 08-AUG-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/474,661  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-067339  
FILING DATE: 04-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-247  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
US-09-140-719-6

Query Match 4.7%; Score 18; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.6e-10;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGGRLYISRM 322  
DB 1 ADLSGIASGGRLYISRM 18

RESULT 6  
US-10-091-442-6  
Sequence 6, Application US/10091442  
Publication No. US20010026931A1  
GENERAL INFORMATION:  
APPLICANT: TSUJIMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSUKUOKA, No. US20010026931A1u  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. US20010026931A1uhiro  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34

YAMAICHI, Kozo  
YAMAGUCHI, No. US20010026931A1omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/091,442  
FILING DATE: 07-Mar-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/140,719  
FILING DATE: 08-AUG-1998  
APPLICATION NUMBER: US 08/474,661  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/091,028  
FILING DATE: 14-JUL-1993  
APPLICATION NUMBER: JP 4-212305  
FILING DATE: 17-JUL-1992  
APPLICATION NUMBER: JP 6-067339  
FILING DATE: 04-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 001560-247  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-091-442-6

Query Match 4.7%; Score 18; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.6e-10;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ADLSGIASGGRLYISRM 322  
DB 1 ADLSGIASGGRLYISRM 18

RESULT 7  
US-09-140-719-4  
Sequence 4, Application US/09140719  
Patent No. US20010026931A1  
GENERAL INFORMATION:  
APPLICANT: TSUJIMOTO, Masafumi  
APPLICANT: IWASA, Fuyuki  
APPLICANT: TSUKUOKA, No. US20010026931A1u  
APPLICANT: NAKAZATO, Hiroshi  
APPLICANT: MIURA, Kenju  
APPLICANT: ISHIDA, No. US20010026931A1uhiro  
APPLICANT: KURIHARA, Tatsuya  
APPLICANT: YAMAICHI, Kozo  
APPLICANT: YAMAGUCHI, No. US20010026931A1omi  
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Burns, Doane, Swecker & Mathis  
;; STREET: P.O. Box 1404  
;; CITY: Alexandria  
;; STATE: Virginia  
;; COUNTRY: United States  
;; ZIP: 22313-1404  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/140,719  
;; FILING DATE: 08-AUG-1998  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/474,661  
;; FILING DATE: 07-JUN-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/091,028  
;; FILING DATE: 14-JUL-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 4-212305  
;; FILING DATE: 17-JUL-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 6-067339  
;; FILING DATE: 04-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McGowan, Malcolm K.  
;; REGISTRATION NUMBER: 39,300  
;; REFERENCE/DOCKET NUMBER: 001560-247  
;; TELEPHONE: (703) 836-6620  
;; TELEFAX: (703) 836-6620  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 17 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-09-140-719-4

Query Match 4.5%; Score 17; DB 9; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 SYIEVTEGTEATAATG 341  
Db 1 SYIEVTEGTEATAATG 17

RESULT 8  
US-10-091-442-4  
; Sequence 4, Application US/10091442  
; Publication No. US20020164711A1  
; GENERAL INFORMATION:  
; APPLICANT: TSUGIMOTO, Masafumi  
; INVENTOR: IWASA, Fuyuki  
; ADDRESSEE: TSURUOKA, No. US20020164711A1uo  
; CITY: MIURA, Kenju  
; STATE: ISHIDA, No. US20020164711A1uhiro  
; COUNTRY: KORIYAMA, Tatsuya  
; ZIP: YAMAGUCHI, Kozo  
; MEDIUM TYPE: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: 34  
; FILING DATE: ADDRESS: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria

;; STATE: Virginia  
;; COUNTRY: United States  
;; ZIP: 22313-1404  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/091,442  
;; FILING DATE: 07-Mar-2002  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/140,719  
;; FILING DATE: 08-AUG-1998  
;; APPLICATION NUMBER: US 08/474,661  
;; FILING DATE: 07-JUN-1995  
;; APPLICATION NUMBER: US 08/091,028  
;; FILING DATE: 14-JUL-1993  
;; APPLICATION NUMBER: JP 4-212305  
;; FILING DATE: 17-JUL-1992  
;; APPLICATION NUMBER: JP 6-067339  
;; FILING DATE: 04-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McGowan, Malcolm K.  
;; REGISTRATION NUMBER: 39,300  
;; REFERENCE/DOCKET NUMBER: 001560-247  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 836-6620  
;; TELEFAX: (703) 836-2021  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 17 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-091-442-4

Query Match 4.5%; Score 17; DB 14; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 SYIEVTEGTEATAATG 341  
Db 1 SYIEVTEGTEATAATG 17

RESULT 9  
US-09-902-684-4  
; Sequence 4, Application US/09902684  
; Patent No. US20020127840A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR  
; INHIBITOR  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/902,684  
; FILING DATE: 12-Jul-2001

```
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/026,408
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: US 60/024,056
/ FILING DATE: 16-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEFFS, ERIC K.
/ REGISTRATION NUMBER: 36,688
/ REFERENCE/DOCKET NUMBER: 1488.0300002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-371-2600
/ TELEFAX: 202-371-2540
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 415 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-902-684-4

Query Match 3.4%; Score 13; DB 10; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
DB 196 MVLNNAVYFKGW 208
|||||

RESULT 10
US-10-094-944-10
/ Sequence 10, Application US/10094944
/ Publication No. US20030180275A1
/ GENERAL INFORMATION:
/ APPLICANT: Immunex Corporation
/ APPLICANT: Clarke, Howard RG
/ APPLICANT: Duosse, Robert F
/ APPLICANT: Wiley, Steven R
/ TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
/ FILE REFERENCE: 3223-A
/ CURRENT APPLICATION NUMBER: US/10/094,944
/ CURRENT FILING DATE: 2002-03-08
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 10
/ LENGTH: 415
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-094-944-10

Query Match 3.4%; Score 13; DB 12; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
DB 196 MVLNNAVYFKGW 208
|||||

RESULT 11
US-10-017-129-4
/ Sequence 4, Application US/10017128
/ Publication No. US20030124536A1
/ GENERAL INFORMATION:
/ APPLICANT: McCarthy, Jeanette
/ TITLE OF INVENTION: DIAGNOSIS AND TREATMENT FOR VASCULAR DISEASE
/ FILE REFERENCE: MM1-001
/ CURRENT APPLICATION NUMBER: US/10/017,128
/ CURRENT FILING DATE: 2001-12-14
/ PRIOR APPLICATION NUMBER: US 60/306,941

/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/026,408
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: US 60/024,056
/ FILING DATE: 16-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEFFS, ERIC K.
/ REGISTRATION NUMBER: 36,688
/ REFERENCE/DOCKET NUMBER: 1488.0300002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-371-2600
/ TELEFAX: 202-371-2540
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 415 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-902-684-4

Query Match 3.4%; Score 13; DB 10; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
DB 196 MVLNNAVYFKGW 208
|||||

RESULT 12
US-10-106-698-4538
/ Sequence 4538, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: PatentIn Ver. 3.0
/ SEQ ID NO 4538
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-106-698-4538

Query Match 3.4%; Score 13; DB 15; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNNAVYFKGW 181
DB 204 MVLNNAVYFKGW 216
|||||

RESULT 13
US-09-140-719-9
/ Sequence 9, Application US/09140719
/ Patent No. US20010026931A1
/ GENERAL INFORMATION:
/ APPLICANT: TSUJIMOTO, Masafumi
/ APPLICANT: IWASA, Fuyuki
/ APPLICANT: TSUBOYOKA, No. US20010026931A1uo
/ APPLICANT: NAKAZATO, Hiroshi
/ APPLICANT: MIURA, Kenju
/ APPLICANT: ISHIDA, No. US20010026931A1uhiro
/ APPLICANT: KURIHARA, Tatsuya
/ APPLICANT: YAMAICHI, Kojo
/ APPLICANT: YAMAGUCHI, No. US20010026931A1omi
/ TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
```



```

; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-140-719-9

Query Match          2.6%; Score 10; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      204 AVAMMHQERK 213
Db      1 AVAMMHQERK 10

RESULT 14
US-10-091-442-9
; Sequence 9, Application US/10091442
; Publication No. US20020164711A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; IWASA, Fuyuki
; TSURUOKA, No. US20020164711A1uo
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, No. US20020164711A1uhiro
; KURIHARA, Tatsuya
; YAMAICHI, Koza
; YAMAGUCHI, No. US20020164711A1omi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States

```

```

; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/091,442
; FILING DATE: 07-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
; US-10-091-442-9

Query Match          2.6%; Score 10; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      204 AVAMMHQERK 213
Db      1 AVAMMHQERK 10

RESULT 15
US-09-993-180-45
; Sequence 45, Application US/09993180
; Publication No. US20030054445A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN SERPIN SECRETED FROM LY
; TITLE OF INVENTION: CELLS, LSI-01
; FILE REFERENCE: 00051.NP
; CURRENT APPLICATION NUMBER: US/09/993,180
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/248,434
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/257,610
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/282,745
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 45
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-993-180-45

Query Match          2.6%; Score 10; DB 11; Length 16;

```

Best Local Similarity 100.0%; Pred. No. 0.017; Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 331 EGGTATAAT 340  
|||||  
Db 4 EGGTATAAT 13

Search completed: December 12, 2003, 16:32:21  
Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:28:03 ; Search time 17 Seconds  
(without alignments)  
1051.185 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 380

Sequence: 1 MASLAANAFQCNLFREMD.....FLFVIRKDDIILFSGKVSCP 380

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	380	100.0	380	1	SPB7_HUMAN
2	18	4.7	380	1	SPB7_MOUSE
3	13	3.4	415	1	PA12_HUMAN
4	13	3.4	415	1	PA12_MOUSE
5	10	2.6	397	1	GDN_MOUSE
6	10	2.6	397	1	GDN_RAT
7	10	2.6	398	1	GDN_HUMAN
8	9	2.4	214	1	SI21_RAT
9	9	2.4	232	1	OVAX_CHICK
10	9	2.4	378	1	PT16_BOVIN
11	9	2.4	382	1	OVAL_COTJA
12	9	2.4	385	1	OVAL_CHICK
13	9	2.4	388	1	OVAX_CHICK
14	9	2.4	405	1	SB12_HUMAN
15	9	2.4	413	1	CP13_RAT
16	8	2.1	291	1	CD20_MOUSE
17	8	2.1	376	1	PT16_HUMAN
18	8	2.1	391	1	SB13_HUMAN
19	8	2.1	423	1	AACT_HUMAN
20	8	2.1	501	1	YSV1_CAEEL
21	8	2.1	922	1	B3AT_CHICK
22	7	1.8	91	1	MERP_ALCSP
23	7	1.8	91	1	MERP_SALTI
24	7	1.8	100	1	VAM3_HUMAN
25	7	1.8	110	1	CHHL_BOMMO
26	7	1.8	165	1	P2Y4_CRIGR
27	7	1.8	170	1	HPAC_SALDI
28	7	1.8	170	1	HPAC_SALTI
29	7	1.8	170	1	HPAC_SALTY
30	7	1.8	188	1	AX2C_PPARAU
31	7	1.8	240	1	YCXN_PORFYR
32	7	1.8	297	1	CD20_HUMAN
33	7	1.8	324	1	DLH1_CANAL

34	7	1.8	361	1	P2Y4_RAT
35	7	1.8	372	1	YD80_BACHD
36	7	1.8	374	1	SPB8_HUMAN
37	7	1.8	375	1	MASP_RAT
38	7	1.8	376	1	SPB9_HUMAN
39	7	1.8	378	1	PT16_MOUSE
40	7	1.8	379	1	ILEU_HUMAN
41	7	1.8	382	1	NDF2_HUMAN
42	7	1.8	392	1	SB11_HUMAN
43	7	1.8	396	1	ARSI_AQUAR
44	7	1.8	410	1	NEUS_CHICK
45	7	1.8	410	1	NEUS_HUMAN

#### ALIGNMENTS

RESULT 1  
SPB7\_HUMAN  
ID SPB7\_HUMAN STANDARD; PRT; 380 AA.  
AC 075635;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Megsin (TF55) (Serpin B7).  
GN SERPINB7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97326116; PubMed=9182567;  
RA Tsujimoto M., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,  
Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,  
Katayama T., Nakao M., Yamashiro K., Hashino J., Haruyama M., Miura K.,  
Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.,  
RT "Purification, cDNA cloning, and characterization of a new serpin with  
megakaryocyte maturation activity".  
RL J. Biol. Chem. 272:15373-15380(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Mesangial cells;  
RC MEDLINE=98376492; PubMed=9710452;  
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,  
Okubo K., Kurokawa K.;  
RT "A mesangium-predominant gene, megsin, is a new serpin upregulated in  
IGA nephropathy".  
RL J. Clin. Invest. 102:828-836(1998).  
CC -!- FUNCTION: Might function as an inhibitor of lys-specific  
proteases. Might influence the maturation of megakaryocytes via  
its action as a serpin.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.  
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D88575; BAA31232.1; -.  
CC EMBL; AF027866; AAC64506.1; -.  
CC HSP; P05619; IHLB.  
CC Genew; HGNC:13902; SERPINB7.  
CC MIM; 603357; -.  
CC GO; GO:0004868; F:serpin; TAS.  
CC InterPro; IPR000215; Serpin.  
CC Pfam; PF00079; serpin; 1.  
CC SMART; SM00063; SERPIN; 1.

```

DR PROSITE; PS00284; SERPIN; 1.
KW Serpin: Serine protease inhibitor.
FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CFF605 CRC64;

Query Match 100.0%; Score 380; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
Matches 380; Conservative 0; Mismatches 0;

QY 1 MASLAANAERCFNLFFEMDDNGNGNFFSSLSLFAALALVRLGAQDLSIQDKLHV 60
DB 1 MASLAANAERCFNLFFEMDDNGNGNFFSSLSLFAALALVRLGAQDLSIQDKLHV 60
QY 61 NTASGYGSSNSQSGLOSLKRVFSDINASHKDYDLSVINGLFAKVFYGHKDYIECAEK 120
DB 61 NTASGYGSSNSQSGLOSLKRVFSDINASHKDYDLSVINGLFAKVFYGHKDYIECAEK 120
QY 121 LYDAKVERVDFTNHLETRNINKWVENETGKIKVIGEGGISSSAWVNVNAVYFKGK 180
DB 121 LYDAKVERVDFTNHLETRNINKWVENETGKIKVIGEGGISSSAWVNVNAVYFKGK 180
QY 181 WQSAFTKGETINCHFKSPKCGKAVAMMEQERKFNLSVIEDPSMKILELYNGINMYVL 240
DB 181 WQSAFTKGETINCHFKSPKCGKAVAMMEQERKFNLSVIEDPSMKILELYNGINMYVL 240
QY 241 LPENDLSIENKLTQNLMEWNTNPRMTSKYVEVFPQPKIEKNYEMKQYLRALGLKQIF 300
DB 241 LPENDLSIENKLTQNLMEWNTNPRMTSKYVEVFPQPKIEKNYEMKQYLRALGLKQIF 300
QY 301 DESKADLSIASGGLYISRMMEKSYIVIEBEGTEATATGNIIVKQLFQSTLFPADHP 360
DB 301 DESKADLSIASGGLYISRMMEKSYIVIEBEGTEATATGNIIVKQLFQSTLFPADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2
SPB7_MOUSE STANDARD; PRT; 380 AA.
AC Q9D695;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Messin (Serpin B7);
GN SERPINB7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21368006; PubMed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagao N., Inagi R., Kurokawa K.;
RT "Cloning of rodent megin revealed its up-regulation in
RT mesangio proliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=SKIN;
RX MEDLINE=21095660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

```

```

Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: Might function as an inhibitor of Lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; AF105328; AAL6768.1; -.
DR EMBL; AK014524; BAB29410.1; -.
DR HSSP; P05121; 1A7C.
DR MGD; MGI:2151053; Serpinb7.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 43050 MW; C9240272ECFB9CF4 CRC64;

Query Match 4.7%; Score 18; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.ee-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWVLNVNAVYFKGW 181
DB 164 SSSAVWVLNVNAVYFKGW 181

RESULT 3
PAI2_HUMAN STANDARD; PRT; 415 AA.
ID PAI2_HUMAN STANDARD; PRT; 415 AA.
AC P05120; Q96E96;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen activator inhibitor-2 precursor (PAI-2) (Plasental
DE plasminogen activator inhibitor) (Monocyte Arg-serpin) (Urokinase
DE inhibitor).
GN SERPIN2 OR PAI2 OR PLANH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RX MEDLINE=87137674; PubMed=3029122;
RA Ye R.D., Wu T.-Z., Sadler J.E.;
RT "cDNA cloning and expression in Escherichia coli of a plasminogen
RT activator inhibitor from human placenta.";
RL J. Biol. Chem. 262:3718-3725(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88142852; PubMed=3325828;
RA Schleuning W.-D., Medcalf R.L., Hession C., Rothenbuhler R.,
RA Shaw A., Kruithof E.K.O.;

```

RT "Plasminogen activator inhibitor 2: regulation of gene transcription  
 RT during phorbol ester-mediated differentiation of U-937 human  
 RT histiocytic lymphoma cells.";  
 RL Mol. Cell. Biol. 7:4564-4567(1987).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=89174589; PubMed=2494165;  
 RX Ye R.D., Ahern S.M., le Beau M.M., Lebo R.V., Sadler J.B.;  
 RA "Structure of the gene for human plasminogen activator inhibitor-2.  
 RT The nearest mammalian homologue of chicken ovalbumin.";  
 RL J. Biol. Chem. 264:5495-5502(1989).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Monocytes;  
 RX MEDLINE=87252928; PubMed=3496414;  
 RA Webb A.C., Collins K.L., Snyder S.F., Alexander S.J.,  
 RA Rosenwasser L.U., Eddy R.L., Shows T.B., Auron P.E.;  
 RT "Human monocyte Arg-Serpin cDNA. Sequence, chromosomal assignment,  
 RT and homology to plasminogen activator-inhibitor.";  
 RL J. Exp. Med. 166:77-94(1987).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Monocytes;  
 RX MEDLINE=48125032; PubMed=3257578;  
 RA Antalis T.M., Clark M.A., Barres T., Lehrbach P.R., Devire P.L.,  
 RA Schevov G., Goss N.H., Stephens R.W., Tolstoshev P.;  
 RT "Cloning and expression of a cDNA coding for a human monocyte-derived  
 RT plasminogen activator inhibitor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:985-989(1988).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=90152678; PubMed=2303256;  
 RX Samia J.A., Alexander S.J., Horton K.W., Auron P.E., Byers M.G.,  
 RA Shows T.B. Jr., Webb A.C.;  
 RT "Chromosomal organization and localization of the human urokinase  
 RT inhibitor gene: perfect structural conservation with ovalbumin.";  
 RL Genomics 6:159-167(1990).  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RX MEDLINE=22389257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Carninci P., Prange C.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Skalius D.E.,  
 RA Scriverch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
 RN [8]  
 RN SEQUENCE OF 12-17; 103-108 AND 314-321.  
 RX MEDLINE=93162043; PubMed=1296667;  
 RA Rasmussen E.H., van Damme J., Puype M., Gesser B., Celis J.E.,  
 RA Vandekerckhove J.;  
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
 RT protein database of normal human epidermal keratinocytes.";  
 RL Electrophoresis 13:960-969(1992).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=99148112; PubMed=10368272;  
 RA Harrop S.J., Jankova L., Coles M., Jardine D., Whittaker J.S.,

RA Gould A.R., Meister A., King G.C., Mabbutt B.C., Curmi P.M.G.;  
 RT "The crystal structure of plasminogen activator inhibitor 2 at 2.0-A  
 RT resolution: implications for serpin function.";  
 RL Structure 7:43-54(1999).  
 RN [10]  
 RN VARIANTS ASP-120; HIS-229; LYS-404 AND CYS-413.  
 RX MEDLINE=99318093; PubMed=10391209;  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [11]  
 RN ERRATUM.  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 CC -1- FUNCTION: PAI-2 INHIBITS UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 CC THE MONOCYTE DERIVED PAI-2 IS DISTINCT FROM THE ENDOTHELIAL  
 CC CELL-DERIVED PAI-1.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC OR EXTRACELLULAR.  
 CC -1- PTM: THE SIGNAL SEQUENCE IS NOT CLEAVED.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use, by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J02685; AAA36413.1; -;  
 DR EMBL; M18082; AAA60006.1; -;  
 DR EMBL; Y00630; CAA68666.1; -;  
 DR EMBL; M24657; AAA60348.1; -;  
 DR EMBL; M24651; AAA60348.1; JOINED.  
 DR EMBL; M24652; AAA60348.1; JOINED.  
 DR EMBL; M24653; AAA60348.1; JOINED.  
 DR EMBL; M24654; AAA60348.1; JOINED.  
 DR EMBL; M24655; AAA60348.1; JOINED.  
 DR EMBL; M24656; AAA60348.1; JOINED.  
 DR EMBL; J03603; AAA60004.1; -;  
 DR EMBL; M31551; AAA36797.1; -;  
 DR EMBL; M31547; AAA36797.1; JOINED.  
 DR EMBL; M31548; AAA36797.1; JOINED.  
 DR EMBL; M31549; AAA36797.1; JOINED.  
 DR EMBL; M31550; AAA36797.1; JOINED.  
 DR EMBL; BC012609; AAHL2609.1; -;  
 DR EMBL; A21238; CAA01535.1; -;  
 DR EMBL; A21254; CAA01539.1; -;  
 DR EMBL; A32853; A32853.  
 DR PDB; 1BY7; 24-OCT-99.  
 DR PDB; 1JRR; 18-DEC-02.  
 DR Aarhus/Ghent-2DPAGE; 6314; 1EP.  
 DR Genew; HGNC:8584; SERPINB2.  
 DR MIM; 173390; -;  
 DR GO; GO:0006916; P:anti-apoptosis; TAS.  
 DR InterPro; IPR00215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;  
 KW Glycoprotein; Signal; 3D-structure; Polymorphism.  
 FT SIGNAL 1 ? NOT CLEAVED.  
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT ACT\_SITE 380 381 REACTIVE BOND.

FT VARIANT 120 120 N -> D (IN dBSNP:6098).  
 FT FTIG=VAR\_011743.  
 FT VARIANT 229 229 R -> H (IN dBSNP:6100).  
 FT FTIG=VAR\_014173.  
 FT VARIANT 404 404 N -> K (IN dBSNP:6103).  
 FT FTIG=VAR\_011744.  
 FT VARIANT 413 413 S -> C (IN dBSNP:6104).  
 FT FTIG=VAR\_011745.  
 FT CONFLICT 170 170 N -> Y (IN REF. 7).  
 FT HELIX 4 22  
 FT STRAND 28 30  
 FT HELIX 32 44  
 FT TURN 45 45  
 FT HELIX 48 57  
 FT TURN 58 63  
 FT HELIX 100 102  
 FT HELIX 103 114  
 FT STRAND 123 132  
 FT TURN 133 134  
 FT STRAND 137 137  
 FT HELIX 139 149  
 FT STRAND 154 156  
 FT HELIX 158 176  
 FT TURN 177 179  
 FT TURN 182 183  
 FT TURN 187 188  
 FT TURN 192 193  
 FT STRAND 196 204

Query Match 3.4%; Score 13; DB 1; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNAVYFKGW 181  
 |||||  
 DB 196 MVLNAVYFKGW 208

## RESULT 4

PA12\_MOUSE  
 ID PA12\_MOUSE STANDARD; PRT; 415 AA.  
 AC P12389; Q35687; Q9QWP6; Q9QWP7; Q9QWP8; Q9QWP9; Q9QWQ; Q9QWZ5;  
 AC Q9QWZ6;  
 DT 01-OCT-1989 (rel. 12, Created)  
 DT 01-OCT-1989 (rel. 12, last sequence update)  
 DT 28-FEB-2003 (rel. 41, last annotation update)  
 DE Plasminogen activator inhibitor-2, macrophage (PAI-2).  
 GN SERPINE2 OR PAI2 OR PLAH2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE OF 1-27 FROM N.A.  
 RC STRAIN=AKR; TISSUE=Myeloid;  
 RX MEDLINE=90059920; PubMed=2583099;  
 RA Belin D., Wohlwend A., Schleuning W.-D., Kruthof E.K.O.,  
 RA Vassalli J.-D.,  
 RT "Facultative polypeptide translocation allows a single mRNA to encode  
 the secreted and cytosolic forms of plasminogen activators inhibitor  
 2.";  
 RT 2.;  
 RL EMBO J. 8:3287-3294(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AKR; TISSUE=Myeloid;  
 RA Belin D.;  
 RN Submitted (SEP-1989) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANTS GUJ-239; TYR-258; ILB-259 AND  
 RP ABG-284.  
 RC STRAIN=129; TISSUE=Liver;  
 RA Belin D., Tapparel C., Sappino N., Silva F.;  
 RT "Sequence and tissue-specific expression of the murine PAI-2 gene.";

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-34; 54-74; 89-115; 124-156; 163-237 AND 265-302 FROM  
 RP N.A.  
 RC STRAIN=129/Sv;  
 RX MEDLINE=99110951; PubMed=9892694;  
 RA Dougherty K.M., Pearson J.M., Yang A.Y., Westrick R.J., Baker M.S.,  
 RA Ginsburg D.;  
 RT "The plasminogen activator inhibitor-2 gene is not required for normal  
 murine development or survival.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:686-691(1999).  
 CC !- FUNCTION: PAI-2 inhibits urokinase-type plasminogen activator. The  
 CC monocyte derived PAI-2 is distinct from the endothelial cell-  
 CC derived PAI-1. Not required for normal murine development or  
 CC survival.  
 CC !- SUBCELLULAR LOCATION: CYTOPLASMIC OR EXTRACELLULAR.  
 CC !- PTM: THE SIGNAL SEQUENCE IS NOT CLEAVED.  
 CC !- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; X16490; CAA34507.1; -;  
 DR EMBL; AJ000386; CAA04047.1; -;  
 DR EMBL; AJ000384; CAA04047.1; JOINED.  
 DR EMBL; AJ000385; CAA04047.1; JOINED.  
 DR EMBL; AF069683; AAD08915.1; -;  
 DR EMBL; AF069685; AAD08916.1; -;  
 DR EMBL; AF069687; AAD08917.1; -;  
 DR EMBL; AF069684; AAD08916.1; JOINED.  
 DR EMBL; AF069686; AAD08917.1; JOINED.  
 DR EMBL; AF069688; AAD08918.1; -;  
 DR EMBL; AF069689; AAD08918.1; JOINED.  
 DR EMBL; AF069690; AAD08919.1; -;  
 DR EMBL; AF069692; AAD08920.1; -;  
 DR EMBL; AF069691; AAD08920.1; JOINED.  
 DR EMBL; AF069694; AAD08921.1; -;  
 DR EMBL; AF069693; AAD08921.1; JOINED.  
 DR HSP; P05120; I8Y7.  
 DR MGD; MGI:97609; Serpinb2.  
 DR InterPro; IPR000240; Maspin.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR PRINTS; PR00676; MASPIN.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;  
 KW Glycoprotein; Signal; Polymorphism.  
 FT SIGNAL 1; NOT CLEAVED.  
 FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT ACT\_SITE 380 381 REACTIVE BOND.  
 FT VARIANT 239 239 A -> E.  
 FT VARIANT 258 258 H -> Y.  
 FT VARIANT 259 259 T -> I.  
 FT VARIANT 284 284 S -> R.  
 FT CONFLICT 55 55 A -> T (IN REF. 3 AND 4).  
 SQ SEQUENCE 415 AA; 46291 MW; C4FC78C5C3CA7D8B CRC64;

Query Match 3.4%; Score 13; DB 1; Length 415;

Best Local Similarity 100.0%; Pred. No. 2.6e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNAVYFKGW 181

DB 196 MVLNAVYFKGW 208

```

RESULT 5
GDN MOUSE
ID GDN MOUSE STANDARD; PRT; 397 AA.
AC Q07235; Q921T7;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE GLIA derived nexin precursor (GDN) (Protease nexin I) (PN-1) (Serine
DE protease-inhibitor-4).
GN SERPINE2 OR PI7 OR PNI OR SPI4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Belin D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Facey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Buterfield Y.S.N., Jones S.J.M., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE OF 1-366 FROM N.A.
RC STAIN-BAL3/C;
RX MEDLINE=93259128; PubMed=8491179;
RA Vassalli J.-D., Huarte J., Bosco D., Sappino A.P., Sappino N.,
RA Vellardi A., Wohlwend A., Erno H., Monard D., Belin D.;
RT "Protease-nexin I as an androgen-dependent secretory product of the
RT murine seminal vesicle."
RL EMBO J. 12:1871-1898 (1993).
CC -!- FUNCTION: THIS GLYCOPROTEIN PROMOTES NEURITE EXTENSION AND IS A
CC SERINE PROTEASE INHIBITOR WITH ACTIVITY TOWARD THROMBIN, TRYPSIN,
CC AND UROKINASE. BINDS HEPARIN.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN SEMINAL VESICLES.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X70236; CAA49777.1; -
CC EXBL; BC010675; AAH10675.1; -
CC EXBL; X70946; CA850285.1; -
CC PIR; I48717; I48717.
CC HSP; P05121; IATC.
CC MGD; MGI:101780; Serpine2.
CC
CC Query Match 2.6%; Score 10; DB 1; Length 397;
CC Best Local Similarity 100.0%; Pred. No. 0.033;
CC Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 170 VLVNAVYFKG 179
Db 182 VLVNAVYFKG 191

RESULT 7
GDN_HUMAN
ID GDN_HUMAN STANDARD; PRT; 398 AA.
AC P07093;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Gla derived nexin precursor (GDN) (Protease nexin I) (FN-1)
DE (Protease inhibitor 7).
GN SERPINE2 OR P17 OR PNI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88107544; PubMed=3427015;
RA Sommer J., Gloor S.M., Rovelli G.F., Hofsteenge J., Nick H.,
RA Meier R., Monard D.;
RT "cDNA sequence coding for a rat glia-derived nexin and its homology
RT to members of the serpin superfamily.";
RL Biochemistry 26:6407-6410(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87051740; PubMed=2877744;
RA Gloor S.M., Odink K., Guenther J., Nick H., Monard D.;
RT "A glia-derived neurite promoting factor with protease inhibitory
RT activity belongs to the protease nexins.";
RL Cell 47:687-693(1986).
RN [3]
RP SEQUENCE FROM N.A.
RA McGrogan M., Kennedy J., Li M.P., Hsu C., Scott R.W., Simonsen C.C.,
RA Baker J.B.;
RT "Molecular cloning and expression of two forms of human protease
RT nexin I.";
RL Biotechnology 6:172-177(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC TIGR=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.N., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 20-47, AND CHARACTERIZATION.
RX MEDLINE=85207723; PubMed=3997857;
RA Scott R.W., Bergman B.L., Bajpai A., Herish R.T., Rodriguez H.,
RA Jones B.N., Barreda C., Watts S., Baker J.B.;
RT "Protease nexin. Properties and a modified purification procedure.";
RL J. Biol. Chem. 260:7029-7034(1985).
CC -!- FUNCTION: THIS GLYCOPROTEIN PROMOTES NEURITE EXTENSION AND IS A
CC SERINE PROTEASE INHIBITOR WITH ACTIVITY TOWARD THROMBIN, TRYPSIN,
CC AND UROKINASE. BINDS HEPARIN.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- MISCELLANEOUS: TWO VARIANTS ARE DERIVED FROM TWO DIFFERENT
CC SUBCLONES.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17783; AAA35883.1; -.
CC EMBL; BC015663; AAH15663.1; -.
CC PIR; A37274; A37274.
CC HSP; P05121; IA7C.
CC Genew; HGNC:8951; SERPINE2.
CC MIM; 177010; -.
CC GO; GO:0005576; Cextracellular; NAS.
CC GO; GO:0008201; F.heparin binding activity; NAS.
CC GO; GO:0004867; F.serine protease inhibitor activity; NAS.
CC GO; GO:0003034; P.regulation of cell migration; NAS.
CC GO; GO:0030162; P.regulation of proteolysis and peptidolysis; NAS.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; SERPIN; 1.
CC PROSITE; PS00284; SERPIN; 1.
CC Serine protease inhibitor; Serpin; Heparin-binding; Neurone;
CC Glycoprotein; Signal.
CC SIGNAL 1 19
CC CHAIN 20 398 GLIA DERIVED NEXIN.
CC FT CARBOHYD 118 118 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 159 159 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT ACT_SITE 365 365 REACTIVE BOND (POTENTIAL).
CC FT VARIANT 329 330 TG -> R.
CC FT FTIG=VAR 007098.
CC FT CONFLICT 261 261 S -> E (IN REF. 2).
CC SEQUENCE 398 AA; 44002 MW; 2A165604E2CBE59 CRC64;
Query Match 2.6%; Score 10; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 170 VLVNAVYFKG 179
Db 182 VLVNAVYFKG 191

RESULT 8
SI21_RAT
ID SI21_RAT STANDARD; PRT; 214 AA.
AC P09005;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DE Serine protease inhibitor 2.1 (SPI-2.1) (Fragment).
DE Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87144617; PubMed=3493437;
RA Hill R.E., Hastie N.D.;
RT "Accelerated evolution in the reactive centre regions of serine
RT protease inhibitors.";
RL Nature 326:96-99(1987).

```



CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X13149; CAA31547.1; -.  
 CC HSSP; P01011; IAS4.  
 CC InterPro; IPR00215; Serpin.  
 CC Pfam; PF00079; serpin; 1.  
 CC SMART; SM00093; SERPIN; 1.  
 CC PROSITE; PS00284; SERPIN; 1.  
 CC Serpin; Serine protease inhibitor.  
 CC NON\_TER 1  
 CC FT ACT SITE 180 181 REACTIVE BOND (BY SIMILARITY).  
 CC SQ SEQUENCE 214 AA; 24218 MW; B1F4CEC59425BAAF CRC64;  
 CC  
 CC Query Match 2.4%; Score 9; DB 1; Length 214;  
 CC Best Local Similarity 100.0%; Pred. No. 0.21;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 333 GTETATATG 341  
 CC DB 166 GTETATATG 174  
 CC  
 CC RESULT 9  
 CC OVAX CHICK  
 CC ID OVAX CHICK STANDARD; PRT; 232 AA.  
 CC AC P01073;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 CC DT 01-APR-1993 (Rel. 25, Last annotation update)  
 CC DE Gene X protein (Ovalbumin-related) (Fragment).  
 CC GN X.  
 CC OS Gallus gallus (Chicken).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Gallinae;  
 CC Gallus.  
 CC NCBI\_TaxID=9031;  
 CC [1]\_TaxID=9031;  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=8102623; PubMed=7418002;  
 CC RA Heilig R., Perrin F., Gannon F., Mandel J.L., Chambon P.;  
 CC RT "The ovalbumin gene family: structure of the X gene and evolution of  
 CC duplicated split genes";  
 CC RL Cell 20:625-637(1980).  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J00920; AAA68881.1; -.  
 CC EMBL; J00918; AAA68881.1; JOINED.  
 CC EMBL; J00919; AAA68881.1; JOINED.  
 CC EMBL; V00385; CAA23683.1; -.  
 CC EMBL; V00386; CAA23684.1; -.  
 CC EMBL; V00387; CAA23685.1; -.  
 CC PIR; A01243; DXCH.  
 CC HSSP; P01012; LOVA.  
 CC InterPro; IPR00215; Serpin.  
 CC Pfam; PF00079; serpin; 1.  
 CC SMART; SM00093; SERPIN; 1.  
 CC PROSITE; PS00284; SERPIN; 1.

KW Serpin.  
 FT NON\_TER 1  
 SQ SEQUENCE 232 AA; 26291 MW; 6B5B86EC4D3B9195 CRC64;  
 CC  
 CC Query Match 2.4%; Score 9; DB 1; Length 232;  
 CC Best Local Similarity 100.0%; Pred. No. 0.22;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 355 FRADHPFLF 363  
 CC DB 205 FRADHPFLF 213  
 CC  
 CC RESULT 10  
 CC PT16 BOVIN  
 CC ID PT16 BOVIN STANDARD; PRT; 378 AA.  
 CC AC 002739;  
 CC DT 30-MAY-2000 (Rel. 39, Created)  
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 CC DE Serine proteinase inhibitor B-43.  
 CC OS Bos taurus (Bovine).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 CC NCBI\_TaxID=9913;  
 CC [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=Brain;  
 CC RX MEDLINE=97165993; PubMed=9013786;  
 CC RA Nakaya N., Nishibori M., Kawabata M., Saeki K.;  
 CC RT "Cloning of a serine proteinase inhibitor from bovine brain:  
 CC expression in the brain and characterization of its target  
 CC proteinases";  
 CC RL Brain Res. Mol. Brain Res. 42:293-300(1996).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES  
 CC PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- TISSUE SPECIFICITY: BRAIN.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D55670; BAA19875.1; -.  
 CC HSSP; P05120; IBY7.  
 CC InterPro; IPR00215; Serpin.  
 CC Pfam; PF00079; serpin; 1.  
 CC SMART; SM00093; SERPIN; 1.  
 CC PROSITE; PS00284; SERPIN; 1.  
 CC Serpin; Serine protease inhibitor.  
 CC ACT SITE 343 344 REACTIVE BOND (BY SIMILARITY).  
 CC SQ SEQUENCE 378 AA; 42560 MW; 664F499CCFCE263A CRC64;  
 CC  
 CC Query Match 2.4%; Score 9; DB 1; Length 378;  
 CC Best Local Similarity 100.0%; Pred. No. 0.34;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 328 EVTEEGTEA 336  
 CC DB 324 EVTEEGTEA 332  
 CC  
 CC RESULT 11  
 CC OVAL COTUA  
 CC ID OVAL COTUA STANDARD; PRT; 382 AA.  
 CC AC P19164;  
 CC DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Ovalbumin.  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 OX NCBI\_TaxID=93334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oviduct;  
 RX MEDLINE=9101650; PubMed=2216734;  
 RA Mucha J., Klaudity J., Klaudinyova V., Hames J., Simuth J.;  
 RT "The sequence of Japanese quail ovalbumin cDNA."  
 RL Nucleic Acids Res. 18:5553-5553(1990).  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X53964; CAA37916.1; -.  
 DR FIR; S11433; S11433.  
 DR HSSP; P01012; LOVA.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF000079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Glycoprotein.  
 FT INIT MET 0 0  
 FT DISULFID 73 120 BY SIMILARITY.  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT ACT SITE 352 353 REACTIVE BOND HOMOLOG.  
 SQ SEQUENCE 382 AA; 42108 MW; A4BE59A1BA5F316 CRC64;  
  
 Query Match 2.4%; Score 9; DB 1; Length 382;  
 Best Local Similarity 100.0%; Pred. No. 0.35;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 355 FRADHPFLF 363  
 |||||  
 DB 355 FRADHPFLF 363  
  
 RESULT 12  
 OVAL\_CHICK STANDARD; PRT; 385 AA.  
 AC P01012;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82069038; PubMed=6272839;  
 RA Woo S.L.C., Beattie W.G., Catterall J.F., Dugaiczky A., Staden R.,  
 RA Brownlee G.G., O'Malley B.W.;  
 RT "Complete nucleotide sequence of the chicken chromosomal ovalbumin  
 RT gene and its biological significance."  
 RL Biochemistry 20:6437-6446(1981).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=78199842; PubMed=661981;

RA McReynolds L., O'Malley B.W., Nisbet A.D., Fothergill J.E., Givol D.,  
 RA Fields S., Robertson M., Brownlee G.G.;  
 RT "Sequence of chicken ovalbumin mRNA."  
 RL Nature 273:723-728(1978).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=79010682; PubMed=692731;  
 RA Catterall J.F., O'Malley B.W., Robertson M.A., Staden R.,  
 RA Tanaka Y., Brownlee G.G.;  
 RT "Nucleotide sequence homology at 12 intron-exon junctions in the  
 RT chick ovalbumin gene."  
 RL Nature 275:510-513(1978).  
 RN [4]  
 RP SEQUENCE OF 1-35.  
 RX MEDLINE=78116057; PubMed=272676;  
 RA Palmier R.D., Gagnon J., Walsh K.A.;  
 RT "Ovalbumin: a secreted protein without a transient hydrophobic leader  
 RT sequence."  
 RL Proc. Natl. Acad. Sci. U.S.A. 75:94-98(1978).  
 RN [5]  
 RP SEQUENCE OF 1-16.  
 RX MEDLINE=79186958; PubMed=751625;  
 RA Thompson E.O.P., Fisher W.K.;  
 RT "A correction and extension of the acetylated amino terminal sequence  
 RT of ovalbumin."  
 RL Aust. J. Biol. Sci. 31:443-446(1978).  
 RN [6]  
 RP SEQUENCE OF 5-16; 29-35; 60-78; 115-123; 366-373 AND 379-385.  
 RX MEDLINE=79186957; PubMed=751624;  
 RA Thompson E.O.P., Fisher W.K.;  
 RT "Amino acid sequences containing half-cystine residues in ovalbumin."  
 RL Aust. J. Biol. Sci. 31:433-442(1978).  
 RN [7]  
 RP N-TERMINUS ACETYLATION, AND PHOSPHORYLATION OF SER-68; SER-236 AND  
 RP SER-240.  
 RX MEDLINE=22056091; PubMed=12060738;  
 RA Maccoss M.J., McDonald W.H., Saraf A., Sadygov R., Clark J.M.,  
 RA Tascio J.J., Gould K.L., Wolters D., Washburn M., Weiss A., Clark J.I.,  
 RA Yates J.R., III;  
 RT "Shotgun identification of protein modifications from protein  
 RT complexes and lens tissue."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:7900-7905(2002).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).  
 RX MEDLINE=90370102; PubMed=2395463;  
 RA Stein P.E., Leslie A.G.W., Finch J.T., Turnell W.G., McLaughlin P.J.,  
 RA Carrell R.W.;  
 RT "Crystal structure of ovalbumin as a model for the reactive centre of  
 RT serpins."  
 RL Nature 347:99-102(1990).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).  
 RX MEDLINE=92046044; PubMed=1942038;  
 RA Stein P.E., Leslie A.G.W., Finch J.T., Carrell R.W.;  
 RT "Crystal structure of uncleaved ovalbumin at 1.95-A resolution."  
 RL J. Mol. Biol. 221:941-959(1991).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=90278960; PubMed=2352279;  
 RA Wright H.T., Qian H.X., Huber R.;  
 RT "Crystal structure of plakalbumin, a proteolytically nicked form of  
 RT ovalbumin. Its relationship to the structure of cleaved alpha-1-  
 RT proteinase inhibitor."  
 RL J. Mol. Biol. 213:513-528(1990).  
 RN [11]  
 RP REVIEW.  
 RX MEDLINE=21312433; PubMed=11419711;  
 RA Huntington J.A., Stein P.E.;  
 RT "Structure and properties of ovalbumin."  
 RL J. Chromatogr. B 756:189-198(2001).  
 CC -!- FUNCTION: Not known.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Major protein of egg white.

CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
CC -!- DATABASE: NAME=Worthington enzyme manual;  
CC WWW="http://www.worthington-biochem.com/manual/OA.html".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J00895; AAB59956.1; -  
CC EMBL; V00438; CAA23716.1; -  
CC EMBL; V00383; CAA23682.1; -  
CC EMBL; M34352; AAA48998.1; -  
CC EMBL; M34346; AAA48998.1; JOINED.  
CC EMBL; M34347; AAA48998.1; JOINED.  
CC EMBL; M34348; AAA48998.1; JOINED.  
CC EMBL; M34349; AAA48998.1; JOINED.  
CC EMBL; M34350; AAA48998.1; JOINED.  
CC EMBL; M34351; AAA48998.1; JOINED.  
CC PIR; A90455; OACH.  
CC PDB; IOVA; 15-JUL-92.  
CC PDB; JUTI; 18-DEC-02.  
CC PDB; 1VAC; 20-JUN-96.  
CC GlycoSuiteDB, P01012, -  
CC InterPro; IPR000215; Serpin.  
CC Pfam; PF000079; serpin; 1.  
CC SMART; SMO00093; SERPIN; 1.  
CC PROSITE; PS00284; SERPIN; 1.  
KW Serpin; Acetylation; Phosphorylation; Glycoprotein; 3D-structure;  
KW Allersen.  
PT INIT MET 0 0  
FT MOD\_RES 1 1  
FT MOD\_RES 68 68  
FT DISULFID 73 120  
FT MOD\_RES 236 236  
FT MOD\_RES 240 240  
FT CARBOHYD 292 292  
FT MOD\_RES 344 344  
FT ACT\_SITE 352 353  
FT VARIANT 311 311 N -> D (IN A MINOR COMPONENT).  
FT CONFLICT 187 187 A -> T (IN REF. 2).  
FT HELIX 3 21  
FT TURN 23 24  
FT STRAND 27 29  
FT HELIX 31 43  
FT TURN 44 44  
FT HELIX 47 57  
FT TURN 58 58  
FT TURN 60 61  
FT TURN 63 64  
FT HELIX 67 70  
FT TURN 71 75  
FT TURN 79 80  
FT HELIX 81 91  
FT STRAND 97 108  
FT TURN 109 110  
FT STRAND 113 113  
FT HELIX 115 124  
FT STRAND 129 132  
FT TURN 135 137  
FT HELIX 138 152  
FT TURN 153 155  
FT TURN 163 164  
FT TURN 168 169  
FT STRAND 172 182  
FT STRAND 184 186  
FT HELIX 190 192  
FT STRAND 194 199  
FT STRAND 205 222  
FT TURN 223 226

FT STRAND 227 234  
FT TURN 235 236  
FT STRAND 239 246  
FT TURN 249 251  
FT HELIX 252 258  
FT TURN 261 267  
FT TURN 268 268  
FT TURN 270 272  
FT STRAND 274 283  
FT STRAND 285 292  
FT HELIX 293 299  
FT TURN 300 301  
FT HELIX 304 306  
FT TURN 308 309  
FT TURN 313 315  
FT STRAND 316 316  
FT TURN 319 320  
FT STRAND 325 334  
FT STRAND 338 340  
FT STRAND 357 359  
FT STRAND 364 370  
FT TURN 371 373  
FT STRAND 376 382  
SQ SEQUENCE 385 AA; 42750 MW; 52B4339642388FE3 CRC64;  
Query Match 2.4%; Score 9; DB 1; Length 385;  
Best Local Similarity 100.0%; Pred.No. 0.35;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 355 FRADHPFLP 363  
DB 358 FRADHPFLP 366  
|||||  
RESULT 13  
ID OVAY\_CHICK STANDARD; PRT; 388 AA.  
AC P01014;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE Gene Y protein (Ovalbumin-related).  
GN Y.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83014329; PubMed=7122240;  
RA Heilig R., Muraskowsky R., Kloeffer C., Mandel J.L.;  
RT "The ovalbumin gene family: complete sequence and structure of the Y  
RT gene."  
RL Nucleic Acids Res. 10:4363-4382(1982).  
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J00922; AAA68882.1; -  
CC PIR; A01244; DYCH.  
CC HSP; P01012; IOVA.  
CC InterPro; IPR000215; Serpin.  
CC Pfam; PF00079; serpin; 1.  
CC SMART; SMO0093; SERPIN; 1.  
CC PROSITE; PS00284; SERPIN; 1.  
KW Serpin.

```
FT ACT_SITE 353 354 REACTIVE BOND HOMOLOG.
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 388 AA; 43772 MW; 2AF17B715D7F461 CRC64;

Query Match 2.4%; Score 9; DB 1; Length 388;
Best Local Similarity 100.0%; Pred.No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFJF 363
Db 361 FRADHPFJF 369

RESULT 14
SB12 HUMAN
ID SB12 HUMAN STANDARD; PRT; 405 AA.
AC Q96P63;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Serpin B12.
GN SERPINB12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=21623618; PubMed=11604408;
RA Askew Y.S., Pak S.C., Luke C.J., Askew D.J., Cataltepe S., Mills D.R.,
RA Kato H., Lehoczy J., Dewar K., Birren B., Silverman G.A.;
RT "SERPINB12 is a novel member of the human ov-serpin family that is
RT widely expressed and inhibits trypsin-like serine proteinases.";
RL J. Biol. Chem. 276:49320-49330(2001).
CC -!- FUNCTION: Inhibits trypsin and plasmin, but not thrombin,
CC coagulation factor Xa, or urokinase-type plasminogen activator.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -!- TISSUE SPECIFICITY: Expressed in many tissues, including brain,
CC bone marrow, lymph node, heart, lung, liver, pancreas, testis,
CC ovary, and intestine.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF411191; AAL05571.1; -.
CC Genew; HGNC:14220; SERPINB12.
CC GO; GO:0005737; C:cytoplasm; NAS.
CC GO; GO:0019899; F:enzyme binding activity; NAS.
CC GO; GO:0004868; F:serpin; IMP.
CC GO; GO:0030304; F:trypsin inhibitor activity; IMP.
CC GO; GO:0042177; P:negative regulation of protein catabolism; NAS.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; Serpin.1
CC SMART; SM00093; SERPIN.1.
CC PROSITE; PS00284; SERPIN.1.
KW Serpin; Serine protease inhibitor.
FT ACT_SITE 370 371 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 405 AA; 46276 MW; FFEL2D4C9B7F3BFA CRC64;

Query Match 2.4%; Score 9; DB 1; Length 405;
Best Local Similarity 100.0%; Pred.No. 0.37;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNAVYFK 178
```

```
Db 186 VLNAVYFK 194

RESULT 15
CPI3_RAT
ID CPI3_RAT STANDARD; PRT; 413 AA.
AC P05544;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Contrapsin-like protease inhibitor 3 precursor (CPI-23) (Serine
DE protease inhibitor 1) (SPI-1).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91324305; PubMed=1864837;
RA Ohkubo K., Ogata S., Misumi Y., Takami N., Ikebara Y.;
RT "Molecular cloning and characterization of rat contrapsin-like
RT protease inhibitor and related proteins.";
RL J. Biochem. 109:243-250(1991).
RN [2]
RP SEQUENCE OF 11-413 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90306038; PubMed=1694763;
RA Pages G., Rouayrenc J.F., le Cam G., Mariller M., le Cam A.;
RT "Molecular characterization of three rat liver serine-protease
RT inhibitors affected by inflammation and hypophysectomy. Protein and
RT mRNA analysis and cDNA cloning.";
RL Eur. J. Biochem. 190:385-391(1990).
RN [3]
RP SEQUENCE OF 82-403 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=87166046; PubMed=3494016;
RA Yoon J.-B., Towle H.C., Seelig S.;
RT "Growth hormone induces two mRNA species of the serine protease
RT inhibitor gene family in rat liver.";
RL J. Biol. Chem. 262:4284-4289(1987).
CC -!- FUNCTION: INHIBITS TRYPSIN, BUT NOT CHYMOTRYPSIN OR
CC ELASTASE.
CC -!- TISSUE SPECIFICITY: LIVER.
CC -!- INDUCTION: BY GROWTH HORMONE.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-11 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D00752; BAA00649.1; -.
CC EMBL; X16357; CAA34406.1; -.
CC EMBL; M15917; AAA42172.1; -.
CC FIR; S08102; S08102.
CC HSP; P01011; ZACH.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; SERPIN.1.
CC PROSITE; PS00284; SERPIN.1.
KW Serpin; serine protease inhibitor; Glycoprotein; Signal.
FT SIGNAL 1 28
FT CHAIN 29 413 CONTRAPSIN-LIKE PROTEASE INHIBITOR 3.
FT ACT_SITE 379 380 REACTIVE BOND (BY SIMILARITY).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. .) (POTENTIAL).
```

FT CARBOHYD 267 267 N-LINKED (GLNAC... ) (POTENTIAL);  
 FT CONFLICT 245 245 V -> L (IN REF. 3);  
 SQ SEQUENCE 413 AA; 46277 MW; 3AACPAPACAAA9BE7 CRC64;  
 Query Match 2.4%; Score 9; DB 1; Length 413;  
 Best Local Similarity 100.0%; Pred. No. 0.37;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 333 GTEATATG 341  
 Db 365 GTEATATG 373  
 |||||

Search completed: December 12, 2003, 16:33:54  
 Job time : 24 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:31:43 ; Search time 21 Seconds  
(without alignments)  
1740.195 Million cell updates/sec

Title: US-09-508-997A-2  
Perfect score: 380  
Sequence: 1 MASLAAANAEFCNLFREMD.....FLFVIRKDDIIILFSCKVSCP 380

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*

1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	3.4	415	2 S20047	plasminogen activa
2	13	3.4	415	2 A32853	plasminogen activa
3	10	2.6	397	2 T30945	hypothetical prote
4	10	2.6	397	2 I48717	proteinase inhibit
5	10	2.6	397	2 B27496	proteinase inhibit
6	10	2.6	398	2 A37274	Glia-derived nexin
7	9	2.4	213	2 A26423	serine proteinase
8	9	2.4	232	1 DXCH	ovalbumin-related
9	9	2.4	248	2 AC1919	hypothetical prote
10	9	2.4	383	2 S11433	ovalbumin - Japan
11	9	2.4	386	1 OACH	ovalbumin [validat
12	9	2.4	388	1 DYCH	ovalbumin-related
13	9	2.4	403	2 S08102	serine proteinase
14	8	2.1	121	2 T31297	hypothetical prote
15	8	2.1	262	2 B83277	hypothetical prote
16	8	2.1	291	2 A30558	B-cell surface ant
17	8	2.1	341	2 H97106	probable permease
18	8	2.1	376	1 A48681	placental thrombin
19	8	2.1	384	2 E97168	probable CDP-4-ret
20	8	2.1	391	2 UC7118	headpin serine pro
21	8	2.1	433	1 ITHUC	alpha-1-antichymot
22	8	2.1	612	2 T35430	probable long-chain
23	8	2.1	613	2 T16885	hypothetical prote
24	8	2.1	626	2 C87672	peptide transpore
25	8	2.1	922	2 A30816	band 3 anion trans
26	8	2.1	1179	2 F71190	probable chromosom
27	7	1.8	19	2 S43652	serpin WSZC1 - whe
28	7	1.8	82	2 AB2351	hypothetical prote
29	7	1.8	121	2 B84770	probable serpin [i

30 7 1.8 123 2 T00549 60S ribosomal prot  
31 7 1.8 152 2 AE2333 hypothetrical prote  
32 7 1.8 157 2 A61017 phosphopyruvate hy  
33 7 1.8 157 2 C71060 hypothetrical prote  
34 7 1.8 167 2 AP2720 hypothetrical prote  
35 7 1.8 170 2 A06030 4-hydroxyphenylace  
36 7 1.8 188 2 T10859 auxin-induced prot  
37 7 1.8 198 2 AC1364 protein gp51 (Bact  
38 7 1.8 220 2 D83495 conserved hypothet  
39 7 1.8 240 2 S73285 hypothetrical prote  
40 7 1.8 242 2 C71006 hypothetrical prote  
41 7 1.8 242 2 T32122 hypothetrical prote  
42 7 1.8 254 2 C82868 plasmid replicatio  
43 7 1.8 261 2 T20264 hypothetrical prote  
44 7 1.8 266 2 C75613 probable molybdenu  
45 7 1.8 270 2 B69781 multidrug-efflux t

## ALIGNMENTS

### RESULT 1

S20047

Plasminogen activator inhibitor 2 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 16-Jul-1999

C/Accession: S20047; S09616

R/Bein, D.

submitted to the EMBL Data Library, September 1989

A/Reference number: S20047

A/Accession: S20047

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-415 <BEL>

A/Cross-references: EMBL:X16490; NID:953589; PIDN:CAA34507.1; PID:953590

R/Bein, D.; Wohlwend, A.; Schleuning, W.D.; Kruthof, E.K.O.; Vassalli, J.D.

EMBO J. 8, 3287-3294, 1989

A/Title: Facultative polypeptide translocation allows a single mRNA to encode the se

A/Reference number: S09615; MUID:90059920; PMID:2583099

A/Accession: S09616

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-27 <BEL2>

A/Cross-references: EMBL:X16490

C/Superfamily: antithrombin III

Query Match 3.4%; Score 13; DB 2; Length 415;  
Best Local Similarity 100.08%; Pred. No. 6.2e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 MVLNVAVYFKGW 181

Db 196 MVLNVAVYFKGW 208

### RESULT 2

A32853

Plasminogen activator inhibitor 2 precursor - human

N/Alternate names: urokinase inhibitor

C/Species: Homo sapiens (man)

C/Date: 22-Nov-1989 #sequence\_revision 22-Nov-1989 #text\_change 20-Oct-2000

C/Accession: A32853; I34218; A31366; A46543; A29362; A26553; A53815; A25021; S02435;

R/Le, R.D.; Ahern, S.M.; Le Beau, M.M.; Lebo, R.V.; Sadler, J.E.

J. Biol. Chem. 264, 5495-5502, 1989

A/Title: Structure of the gene for human plasminogen activator inhibitor-2. The near

A/Reference number: A32853; MUID:89174589; PMID:2494165

A/Accession: A32853

A/Molecule type: DNA

A/Residues: 1-415 <YER>

A/Cross-references: GB:M24651; GB:M24652; GB:M24653; GB:M24654; GB:M24655; GB:M24656

R/Samia, J.A.; Alexander, S.J.; Horton, K.W.; Byers, M.G.; Shows, T.B.

Genomics 6, 159-167, 1990

A/Title: Chromosomal organization and localization of the human urokinase inhibitor

A;Reference number: I54218; MUID:90152578; PMID:2303256  
 A;Accession: I54218  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-415 <RES>  
 A;Cross-references: GB:M31551; NID:G340152; PIDN:AAA36797.1; PID:G340154  
 R;Anfalls, I.M.; Clark, M.A.; Barnes, T.; Lehrbach, P.R.; Devine, P.L.; Schvezov, G.; G  
 Proc. Natl. Acad. Sci. U.S.A. 85, 985-989, 1988  
 A;Title: Cloning and expression of a cDNA coding for a human monocyte-derived plasminoge  
 A;Reference number: A31366; MUID:88125032; PMID:3257578  
 A;Accession: A31366  
 A;Molecule type: mRNA  
 A;Residues: 1-415 <ANT>  
 A;Cross-references: GB:J03603; NID:G189546; PIDN:AAA60004.1; PID:G189547  
 R;Webb, A.C.; Collins, K.L.; Snyder, S.E.; Alexander, S.J.; Roserwasser, L.J.; Eddy, R.I  
 J. Exp. Med. 166, 77-94, 1997  
 A;Title: Human monocyte Arg-serpin cDNA. Sequence, chromosomal assignment, and homology  
 A;Reference number: A46543; MUID:8725228; PMID:3496414  
 A;Accession: A46543  
 A;Molecule type: mRNA  
 A;Residues: 1-415 <WEB>  
 A;Cross-references: GB:Y00630; NID:G35267; PIDN:CAA68666.1; PID:G35268  
 R;Schleuning, W.D.; Medcalf, R.L.; Hession, C.; Rothenbuehler, R.; Shaw, A.; Knutthof, E  
 Mol. Cell. Biol. 7, 4564-4567, 1987  
 A;Title: Plasminogen activator inhibitor 2: regulation of gene transcription during phor  
 A;Reference number: A29362; MUID:88142852; PMID:3325828  
 A;Accession: A29362  
 A;Molecule type: mRNA  
 A;Residues: 1-119, 'D', 121-403, 'K', 405-412, 'C', 414-415 <SCH>  
 A;Cross-references: GB:M18082; NID:G189562; PIDN:AAA60006.1; PID:G189563  
 R;Ye, R.D.; Wu, T.C.; Sadler, J.E.  
 J. Biol. Chem. 262, 3718-3725, 1987  
 A;Title: cDNA cloning and expression in Escherichia coli of a plasminogen activator inh  
 A;Reference number: A26553; MUID:87137674; PMID:3029122  
 A;Accession: A26553  
 A;Molecule type: mRNA  
 A;Residues: 1-119, 'D', 121-403, 'K', 405-412, 'C', 414-415 <YE2>  
 A;Cross-references: GB:J02895; NID:G189544; PIDN:AAA36413.1; PID:G189545  
 R;Experimental Source: placenta  
 R;Jensen, P.H.; Schueler, E.; Woodrow, G.; Richardson, M.; Goss, N.; Hojrup, P.; Peterse  
 J. Biol. Chem. 269, 15394-15398, 1994  
 A;Title: A unique interhelical insertion in plasminogen activator inhibitor-2 contains  
 A;Reference number: A53815; MUID:94253109; PMID:7910824  
 A;Accession: A53815  
 A;Molecule type: protein  
 A;Residues: 83-67 <JEN>  
 R;Kruithof, E.K.O.; Vassalli, J.D.; Schleuning, W.D.; Mattaliano, R.J.; Bachmann, F.  
 J. Biol. Chem. 261, 11207-11213, 1986  
 A;Title: Purification and characterization of a plasminogen activator inhibitor from the  
 A;Reference number: A25021; MUID:86278222; PMID:3090045  
 A;Accession: A25021  
 A;Molecule type: protein  
 A;Residues: 347-376 <KRU>  
 R;Kiso, U.; Kaudewitz, H.; Henschen, A.; Astedt, B.; Kruithof, E.K.O.; Bachmann, F.  
 FEBS Lett. 230, 51-56, 1988  
 A;Title: Determination of intermediates, products and cleavage site in the reaction betw  
 A;Reference number: S02435; MUID:88167197; PMID:3280346  
 A;Accession: S02435  
 A;Molecule type: protein  
 A;Residues: 377-415 <KIS>  
 R;Kruithof, E.K.; Cousin, E.  
 Biochem. Biophys. Res. Commun. 156, 383-388, 1988  
 A;Title: Plasminogen activator inhibitor 2. Isolation and characterization of the promot  
 A;Reference number: I52229; MUID:89025873; PMID:2845977  
 A;Accession: I52229  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-56 <RE2>  
 A;Cross-references: GB:M23092; NID:G189559; PIDN:AAA60005.1; PID:G189561  
 C;Genetics:

A;Introns: 56/3; 96/3; 139/3; 179/1; 226/3; 281/3  
 C;Superfamily: antithrombin III  
 C;Keywords: acetylated amino end; glycoprotein; serine proteinase inhibitor  
 F;1-22/Domain: signal sequence #status predicted <SIG>  
 F;23-415/Product: plasminogen activator inhibitor 2 #status predicted <MAT>  
 F;23/Modified site: acetylated amino end (Ser) (in mature form) #status predicted  
 F;75,115,339/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F;83,84,86/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of unidentified p  
 F;380-381/Cleavage site: Arg-Thr (plasminogen activator) #status experimental  
 F;380/Inhibitory site: Arg (plasminogen activator) #status experimental  
 Query Match 3.4%; Score 13; DB 2; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 169 MVLVNAVYFKGK 181  
 |||||  
 Db 196 MVLVNAVYFKGK 208  
 |||||  
 RESULT 3  
 T30945  
 hypothetical protein AC7.3 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jun-2000  
 C;Accession: T30945  
 R;Nelson, J.; Hawkins, J.  
 submitted to the EMBL Data Library, June 1999  
 A;Description: The sequence of C. elegans cosmid AC7.  
 A;Reference number: Z20944  
 A;Accession: T30945  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-238 <NEL>  
 A;Cross-references: EMBL:U61957; PIDN:AA03419.1  
 A;Experimental source: strain Bristol N2; clone AC7  
 C;Genetics:  
 A;Map position: IV  
 A;Introns: 10/2; 30/3; 72/3; 159/3  
 A;Note: AC7.3  
 C;Superfamily: Caenorhabditis elegans hypothetical protein AC7.3  
 Query Match 2.6%; Score 10; DB 2; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 0.047;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 146 VENETHGKIK 155  
 |||||  
 Db 92 VENETHGKIK 101  
 |||||  
 RESULT 4  
 I48717  
 proteinase inhibitor nexin I precursor - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Jul-1999  
 C;Accession: I48717; S70772; S35731  
 R;Vassalli, J.D.; Huarte, J.; Bosco, D.; Sappino, A.P.; Velardi, A.; Wol  
 EMBO J. 12, 1871-1878, 1993  
 A;Title: Protease-nexin I as an androgen-dependent secretory product of the murine se  
 A;Reference number: I48717; MUID:93259128; PMID:8491179  
 A;Accession: I48717  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-397 <RES>  
 A;Cross-references: EMBL:X70296; NID:G551064; PIDN:CAA49777.1; PID:G551065  
 A;Accession: S70772  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-86 <VAS>  
 A;Cross-references: EMBL:X70946; NID:G57930; PIDN:CAA50285.1; PID:G57931  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1993;  
 C;Genetics:

A:Gene: FN-1  
C:Superfamily: antithrombin III  
C:Keywords: serine proteinase inhibitor  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-397/Product: proteinase inhibitor nexin I #status predicted <NAT>

Query Match 2.6%; Score 10; DB 2; Length 397;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNVAVYFKG 179  
| | | | | | | | | |  
Db 182 VLNVAVYFKG 191

RESULT 5  
B27496  
proteinase inhibitor nexin I precursor - rat (fragment)  
N:Alternate names: glia-derived nexin (GDN)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text\_change 16-Jul-1999  
C:Accession: B27496; A34538; A42351; B42351; C42351  
R:Sommer, J.; Gloor, S.M.; Rovelli, G.F.; Hofsteenge, J.; Nick, H.; Meier, R.; Monard, D.  
Biochemistry 26, 6407-6410, 1987  
A:Title: cDNA sequence coding for a rat glia-derived nexin and its homology to members  
A:Reference number: A90519; MUID:88107544; PMID:3427015  
A:Accession: B27496  
A:Molecule type: mRNA  
A:Residues: 1-397 <SOM>  
A:Cross-references: GB:M17783; NID:G204283; PIDN:AAA1209.1; PID:G204284  
A>Note: the authors translated the codon TGG for residue 156 as Thr  
R:Nick, H.; Hofsteenge, J.; Shaw, E.; Rovelli, G.; Monard, D.  
Biochemistry 29, 2417-2421, 1990  
A:Title: Functional sites of glia-derived nexin (GDN): importance of the site reacting w  
A:Reference number: A34538; MUID:90248459; PMID:2337608  
A:Accession: A34538  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 364-394 <NIC>  
R:Rovelli, G.; Stone, S.R.; Guidolin, A.; Sommer, J.; Monard, D.  
Biochemistry 31, 3542-3549, 1992  
A:Title: Characterization of the heparin-binding site of glia-derived nexin/protease nex  
A:Reference number: A42351; MUID:92207980; PMID:1554734  
A:Accession: A42351  
A:Molecule type: protein  
A:Residues: 82-96 <ROV1>  
A>Note: sequence extracted from NCBI backbone (NCBIP:93846)  
A:Accession: B42351  
A:Molecule type: protein  
A:Residues: 165-178 <ROV2>  
A>Note: sequence extracted from NCBI backbone (NCBIP:93851)  
A:Accession: C42351  
A:Molecule type: protein  
A:Residues: 317-333 <ROV3>  
A>Note: sequence extracted from NCBI backbone (NCBIP:93856)  
A>Note: peptide sequences were determined from rat cDNA cloned and expressed in yeast  
C:Superfamily: antithrombin III  
C:Keywords: glycoprotein; serine proteinase inhibitor  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-397/Product: proteinase inhibitor nexin I #status predicted <NAT>  
F:159/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.6%; Score 10; DB 2; Length 397;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNVAVYFKG 179  
| | | | | | | | | |  
Db 182 VLNVAVYFKG 191

RESULT 6  
A37274  
serine proteinase inhibitor 2.1 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 19-Nov-1988 #sequence\_revision 12-Jul-1996 #text\_change 12-Jul-1996  
C:Accession: A26423  
R:Hill, R.E.; Hastie, N.D.  
Nature 326, 96-99, 1987  
A:Title: Accelerated evolution in the reactive centre regions of serine protease inhi  
A:Reference number: A26423; MUID:87144617; PMID:3493437  
A:Accession: A26423  
A:Molecule type: mRNA

Glia-derived nexin I precursor, splice form beta - human  
N:Alternate names: glia-derived neurite promoting factor; proteinase inhibitor 7; pm  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text\_change 20-Jun-2000  
C:Accession: A37274; J00010; A27496; A26061; A24051  
R:McGrogan, M.; Kennedy, J.; Li, M.P.; Hsu, C.; Scott, R.W.; Simonsen, C.C.; Baker, A  
BioTechnology 6, 172-177, 1988  
A:Title: Molecular cloning and expression of two forms of human protease nexin I.  
A:Reference number: J00010  
A:Accession: A37274  
A:Molecule type: mRNA  
A:Residues: 1-398 <MCG1>  
A:Experimental source: splice form beta  
A:Accession: J00010  
A:Molecule type: mRNA  
A:Residues: 1-328, 'R', 331-398 <MCG2>  
A:Experimental source: splice form alpha  
R:Sommer, J.; Gloor, S.M.; Rovelli, G.F.; Hofsteenge, J.; Nick, H.; Meier, R.; Monard  
Biochemistry 26, 6407-6410, 1987  
A:Title: cDNA sequence coding for a rat glia-derived nexin and its homology to membe  
A:Reference number: A90519; MUID:88107544; PMID:3427015  
A:Accession: A27496  
A:Molecule type: mRNA  
A:Residues: 1-328, 'R', 331-398 <SOM>  
A:Cross-references: GB:M17783; NID:G183063; PIDN:AAA35883.1; PID:G183064  
R:Gloor, S.; Odink, K.; Guenther, J.; Nick, H.; Monard, D.  
Cell 47, 687-693, 1986  
A:Title: A glia-derived neurite promoting factor with protease inhibitory activity b  
A:Reference number: A26061; MUID:87051740; PMID:2877744  
A:Accession: A26061  
A:Molecule type: protein  
A:Residues: 1-259, 'S', 261-398 <GLO>  
R:Scott, R.W.; Bergman, B.L.; Baijap, A.; Herish, R.T.; Rodriguez, H.; Jones, B.N.; B  
J. Biol. Chem. 260, 7029-7034, 1985  
A:Title: Protease nexin. Properties and a modified purification procedure.  
A:Reference number: A24051; MUID:85207723; PMID:3997857  
A:Accession: A24051  
A:Molecule type: protein  
A:Residues: 20-47 <SCO>  
C:Genetics:  
A:Gene: GDB:PI7; PNI  
A:Cross-references: GDB:378380; OMIM:177010  
A:Map position: 2q33-2q35  
C:Superfamily: antithrombin III  
C:Keywords: alternative splicing; glycoprotein; serine proteinase inhibitor  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-398/Product: glia-derived nexin I, splice form beta #status experimental <NATA>  
F:20-328, 'R', 331-398/Product: glia-derived nexin I, splice form alpha #status predict  
F:118,159,394/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:365/Inhibitory site: Arg (thrombin, urokinase) #status predicted

Query Match 2.6%; Score 10; DB 2; Length 398;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNVAVYFKG 179  
| | | | | | | | | |  
Db 182 VLNVAVYFKG 191

RESULT 7  
A26423  
serine proteinase inhibitor 2.1 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 19-Nov-1988 #sequence\_revision 12-Jul-1996 #text\_change 12-Jul-1996  
C:Accession: A26423  
R:Hill, R.E.; Hastie, N.D.  
Nature 326, 96-99, 1987  
A:Title: Accelerated evolution in the reactive centre regions of serine protease inhi  
A:Reference number: A26423; MUID:87144617; PMID:3493437  
A:Accession: A26423  
A:Molecule type: mRNA



A;Residues: 1-213 <HIL>  
A;Cross-references: GB:X13149  
A;Note: the authors translated the codon CTG for residue 1 as Met, GTG for residue 35 as Met  
C;Superfamily: antithrombin III

Query Match 2.4%; Score 9; DB 2; Length 213;  
Best Local Similarity 100.0%; Pred. No. 0.46;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 333 GTEATAATG 341  
|||||  
DB 165 GTEATAATG 173

RESULT 8  
DXCH  
ovalbumin-related X protein - chicken (fragment)  
C;Species: Gallus gallus (chicken)  
C;Date: 31-Dec-1980 #sequence\_revision 31-Dec-1980 #text\_change 18-Jun-1999  
C;Accession: A01243  
R;Heilig, R.; Perrin, F.; Gannon, F.; Mardel, J.L.; Chambon, P.  
Cell 20, 625-637, 1990  
A;Title: The ovalbumin gene family: structure of the X gene and evolution of duplicated F13/Binding site: carbohydrate (Asn) (covalent) #status predicted  
A;Reference number: A01243; MUID:81022623; PMID:7418002  
A;Contents: exons 6, 7, and 8  
A;Accession: A01243  
A;Molecule type: DNA  
A;Residues: 1-232 <HEI>  
C;Cross-references: GB:J00920; NID:g212695; PIDN:AAA68881.1; PID:g212897  
C;Superfamily: antithrombin III  
C;Keywords: glycoprotein; phosphoprotein; serine proteinase inhibitor  
F13/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F189/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 2.4%; Score 9; DB 1; Length 233;  
Best Local Similarity 100.0%; Pred. No. 0.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFLF 363  
|||||  
DB 205 FRADHPFLF 213

RESULT 9  
AC1919  
hypothetical protein all0902 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C;Accession: AC1919  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shingo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AC1919  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-248 <KUR>  
C;Cross-references: GB:BA000019; PIDN:BA072859.1; PID:g17130248; GSPDB:GN00179  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: all0902

Query Match 2.4%; Score 9; DB 2; Length 248;  
Best Local Similarity 100.0%; Pred. No. 0.53;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 DLSEIENKL 253  
|||||  
DB 236 DLSEIENKL 244

RESULT 10  
S11433  
ovalbumin - Japanese quail  
C;Species: Coturnix coturnix japonica (Japanese quail)  
C;Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
C;Accession: S11433  
R;Nucha, J.; Klaudiny, J.; Klaudinyova, V.; Hanes, J.; Simuth, J.  
Nucleic Acids Res. 18, 5553, 1990  
A;Title: The sequence of Japanese quail ovalbumin cDNA.  
A;Reference number: S11433; MUID:91016850; PMID:2216734  
A;Accession: S11433  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-383 <MUC>  
A;Cross-references: EMBL:X53964; NID:g62643; PIDN:CAA37916.1; PID:g62644  
C;Superfamily: antithrombin III

Query Match 2.4%; Score 9; DB 2; Length 383;  
Best Local Similarity 100.0%; Pred. No. 0.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFLF 363  
|||||  
DB 356 FRADHPFLF 364

RESULT 11  
CACH  
ovalbumin [validated] - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 31-Dec-1979 #sequence\_revision 30-Jun-1993 #text\_change 15-Sep-2000  
C;Accession: A90455; ID:150402; ID:150402; A93197; A93827; A90093; A90092; A61297; A42793;  
R;Woo, S.L.C.; Beattie, W.G.; Catterall, J.F.; Dugaiczky, A.; Staden, R.; Brownlee, G.  
Biochemistry 20, 6437-6446, 1981  
A;Title: Complete nucleotide sequence of the chicken chromosomal ovalbumin gene and its expression in oviduct epithelial cells  
A;Reference number: A90455; MUID:82069038; PMID:6272839  
A;Accession: A90455  
A;Molecule type: DNA  
A;Residues: 1-386 <WOO>  
A;Cross-references: EMBL:V00438; NID:g63719; PIDN:CAA23716.1; PID:g808974  
A;Note: a number of silent polymorphic sites are identified and discussed  
A;Note: Thr-188 is also predicted  
R;Catterall, J.F.; O'Malley, B.W.; Robertson, M.A.; Staden, R.; Tanaka, Y.; Brownlee, N.  
Nature 275, 510-513, 1978  
A;Title: Nucleotide sequence homology at 12 intron-exon junctions in the chick ovalbumin gene  
A;Reference number: ID:150402; MUID:79010682; PMID:692731  
A;Accession: ID:150402  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-386 <CAT>  
A;Cross-references: GB:M34352; NID:g212501; PIDN:AAA48998.1; PID:g212503  
R;Robertson, M.A.; Staden, R.; Tanaka, Y.; Catterall, J.F.; O'Malley, B.W.; Brownlee, N.  
Nature 278, 370-372, 1979  
A;Title: Sequence of three introns in the chick ovalbumin gene.  
A;Reference number: ID:150605; MUID:79135070; PMID:423993  
A;Accession: ID:150605  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-4, 'A', 'F', '120-155 <ROB>  
A;Cross-references: EMBL:V00382; NID:g63051; PIDN:CAA23681.1; PID:g63052  
R;McReynolds, L.; O'Malley, B.W.; Nisbet, A.D.; Pothergill, J.E.; Givol, D.; Fields, N.  
Nature 273, 723-728, 1978  
A;Title: Sequence of chicken ovalbumin mRNA.  
A;Reference number: A93197; MUID:78199842; PMID:661981  
A;Accession: A93197  
A;Molecule type: mRNA  
A;Residues: 1-386 <MCR>  
A;Cross-references: EMBL:V00383; NID:g63053  
A;Note: a minor component has Asp-312  
R;Palmiter, R.D.; Gagnon, J.; Walsh, K.A.  
Proc. Natl. Acad. Sci. U.S.A. 75, 94-98, 1978  
A;Title: Ovalbumin: a secreted protein without a transient hydrophobic leader sequen

A;Reference number: A93827; MUID:78115057; PMID:272676  
A;Accession: A93827  
A;Molecule type: protein  
A;Residues: 2-33, X, 35-36 <PAL>  
R;Thompson, E.O.P.; Fisher, W.K.  
Aust. J. Biol. Sci. 31, 443-446, 1978  
A;Title: A correction and extension of the acetylated amino terminal sequence of ovalbumin  
A;Reference number: A90093; MUID:79186958; PMID:751625  
A;Accession: A90093  
A;Molecule type: protein  
A;Residues: 2-17 <TH1>  
R;Thompson, E.O.P.; Fisher, W.K.  
Aust. J. Biol. Sci. 31, 433-442, 1978  
A;Title: Amino acid sequences containing half-cysteine residues in ovalbumin.  
A;Reference number: A90092; MUID:79186957; PMID:751624  
A;Accession: A90092  
A;Molecule type: protein  
A;Residues: 6-17,30-36;61-79,116-124,367-374,380-386 <TH2>  
R;Tsunasawa, S.; Narita, K.  
J. Biochem. 92, 607-613, 1982  
A;Title: Micro-identification of amino-terminal acetyl amino acids in proteins.  
A;Reference number: A61297; MUID:43056735; PMID:6754709  
A;Accession: A61297  
A;Molecule type: protein  
A;Residues: 2-6 <TSU>  
R;Takahashi, N.; Hirose, M.  
J. Biol. Chem. 267, 11565-11572, 1992  
A;Title: Reversible denaturation of disulfide-reduced ovalbumin and its reoxidation gene  
A;Reference number: A42793; MUID:92283876; PMID:1597494  
A;Accession: A42793  
A;Molecule type: protein  
A;Residues: 60-73, X, 75-85;112-119, 'EX', 122-123 <TAK>  
R;Stein, P.B.; Leslie, A.G.W.  
submitted to the Brookhaven Protein Data Bank, November 1990  
A;Reference number: A50294; PDB:1OXA  
A;Contents: annotation; X-ray crystallography, 1.95 angstroms, residues 2-386  
R;Stein, P.B.; Leslie, A.G.W.; Finch, J.T.; Carrell, R.W.  
J. Mol. Biol. 221, 941-959, 1991  
A;Title: Crystal structure of uncleaved ovalbumin at 1.95 Angstroms resolution.  
A;Reference number: A58761; MUID:92046044; PMID:1942038  
A;Contents: annotation; X-ray crystallography, 1.95 angstroms  
C;Genetics:  
A;Introns: 56/3; 73/3; 116/3; 156/1; 203/3; 255/3  
C;Superfamily: antithrombin III  
C;Keywords: acetylated amino end; glycoprotein; phosphoprotein  
F;2-36/Product: ovalbumin #status experimental <MAR>  
F;2/Modified site: acetylated amino end (Gly) (in mature form) #status experimental  
F;59,345/Binding site: phosphate (Ser) (covalent) #status experimental  
F;74-121/Disulfide bonds: #status experimental  
F;293/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 2.4%; Score 9; DB 1; Length 386;  
Best Local Similarity 100.0%; Pred. No. 0.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFLF 363  
|||||  
Db 359 FRADHPFLF 367

RESULT 12  
DYCH  
ovalbumin-related y protein - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 18-Jun-1999  
C;Accession: A01244  
R;Heilig, R.; Muraskovsky, R.; Kloepper, C.; Mandel, J.L.  
Nucleic Acids Res. 10, 4363-4382, 1982  
A;Title: The ovalbumin gene family: complete sequence and structure of the y gene.  
A;Reference number: A01244; MUID:83014329; PMID:17122240  
A;Accession: A01244  
A;Molecule type: DNA  
A;Residues: 1-388 <HEI>

A;Cross-references: GB:J00922; GB:V00439; NID:g212899; PIDN:AAA68882.1; PID:g212900  
C;Genetics:  
A;Introns: 56/3; 73/3; 116/3; 156/1; 203/3; 255/3  
C;Superfamily: antithrombin III  
C;Keywords: glycoprotein; phosphoprotein; serine proteinase inhibitor  
F;74-121/Disulfide bonds: #status predicted  
F;293/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;345/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 2.4%; Score 9; DB 1; Length 388;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 FRADHPFLF 363  
|||||  
Db 361 FRADHPFLF 369

RESULT 13  
S08102  
serine proteinase inhibitor 1 - rat  
N;Alternate names: growth hormone-induced proteinase inhibitor  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 16-Jul-1999  
C;Accession: S08102; S11318; A29131  
R;le Cam, A.  
submitted to the EMBL Data Library, August 1989  
A;Reference number: S08099  
A;Accession: S08102  
A;Molecule type: mRNA  
A;Residues: 1-403 <LEC>  
A;Cross-references: EMBL:X16357; NID:g57230; PIDN:CAA34406.1; PID:g57231  
R;Pages, G.; Rouayrenc, J.F.; le Cam, G.; Mariller, M.; le Cam, A.  
Eur. J. Biochem. 190, 385-391, 1990  
A;Title: Molecular characterization of three rat liver serine-protease inhibitors af  
A;Reference number: S11318; MUID:90306038; PMID:1694763  
A;Accession: S11318  
A;Status: Preliminary  
A;Molecule type: mRNA  
A;Residues: 1-403 <PAG>  
R;Yoon, J.B.; Towle, H.C.; Seelig, S.  
J. Biol. Chem. 262, 4284-4289, 1987  
A;Title: Growth hormone induces two mRNA species of the serine protease inhibitor ge  
A;Reference number: A92632; MUID:87166046; PMID:3494016  
A;Accession: A29131  
A;Molecule type: mRNA  
A;Residues: 82-234, 'L', 236-403 <YOO>  
A;Cross-references: GB:M15917; GB:J02692; NID:g207041; PIDN:AAA42172.1; PID:g207042  
C;Superfamily: antithrombin III

Query Match 2.4%; Score 9; DB 2; Length 403;  
Best Local Similarity 100.0%; Pred. No. 0.84;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 333 GTTATATG 341  
|||||  
Db 355 GTTATATG 363

RESULT 14  
T31297  
hypothetical protein - Sphingomonas aromaticivorans plasmid pNL1  
C;Species: Sphingomonas aromaticivorans  
C;Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000  
C;Accession: T31297  
R;Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.  
submitted to the EMBL Data Library, July 1998  
A;Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aro  
A;Reference number: Z20992  
A;Accession: T31297  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

A:Residues: 1-121 <ROM>  
A:Cross-references: EMBL:AF079317; NID:G3378261; PID:G3378438; PIDN:AAD04021.1  
C:Genetics:  
A:Genome: plasmid pNL1  
A>Note: pchC

Query Match 2.1%; Score 8; DB 2; Length 121;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 LSLFAALA 40  
| | | | | | | |  
Db 7 LSLFAALA 14

RESULT 15  
B83827  
Hypothetical protein BHI418 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: B83827  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: B83827  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <STO>  
A:Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA005137.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BHI418  
C:Superfamily: Hypothetical protein M50933

Query Match 2.1%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASIAAANA 9  
| | | | | | | |  
Db 111 ASIAAANA 118

Search completed: December 12, 2003, 16:35:17  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:31:13 ; Search time 40 Seconds  
(without alignments)  
2451.500 Million cell updates/sec

Title: US-09-508-997A-2  
Perfect score: 380  
Sequence: 1 MASLAANAEPFCNLFREMD.....FVIRKDDILFSGVSCP 380

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp Unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteria:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	18	4.7	380	11	Q9D695	Q9D695 mus musculus
2	18	4.7	380	11	Q920U5	Q920U5 rattus norv
3	14	3.7	334	12	Q9DHG4	Q9DHG4 yaba-like d
4	11	2.9	385	11	Q8VHQ0	Q8VHQ0 mus musculus
5	11	2.9	388	11	Q92ZG2	Q92ZG2 mus musculus
6	10	2.6	292	6	Q8WNW8	Q8WNW8 sus scrofa
7	10	2.6	359	11	Q9D6A7	Q9D6A7 mus musculus
8	10	2.6	366	11	Q9DIE7	Q9DIE7 mus musculus
9	10	2.6	387	11	Q9DIQ5	Q9DIQ5 mus musculus
10	10	2.6	387	11	Q8BHL1	Q8BHL1 mus musculus
11	10	2.6	387	11	Q8BGB6	Q8BGB6 mus musculus
12	10	2.6	397	6	Q8HZY1	Q8HZY1 bos taurus
13	9	2.4	248	16	Q8VYE8	Q8VYE8 anabaena sp
14	9	2.4	423	11	Q9DFP9	Q9DFP9 mus musculus
15	9	2.4	454	16	Q92K34	Q92K34 rhizobium m
16	9	2.4	615	4	Q8TAP0	Q8TAP0 homo sapien

17	9	2.4	615	4	Q9NUX3	Q9NUX3 homo sapien
18	8	2.1	121	2	Q86005	Q86005 sphingomona
19	8	2.1	195	4	Q96DW8	Q96DW8 homo sapien
20	8	2.1	233	11	Q8BNW4	Q8BNW4 mus musculus
21	8	2.1	259	4	Q8IXH2	Q8IXH2 homo sapien
22	8	2.1	262	16	Q9KD01	Q9KD01 bacillus ha
23	8	2.1	288	5	Q9GZ39	Q9GZ39 trypanosoma
24	8	2.1	341	16	Q97IG4	Q97IG4 clostridium
25	8	2.1	381	5	Q02377	Q02377 manduca sex
26	8	2.1	384	16	Q97H35	Q97H35 clostridium
27	8	2.1	401	16	Q92QU5	Q92QU5 rhizobium m
28	8	2.1	407	4	Q9UNU9	Q9UNU9 homo sapien
29	8	2.1	423	4	Q8N177	Q8N177 homo sapien
30	8	2.1	537	5	Q9VA64	Q9VA64 drosophila
31	8	2.1	612	16	Q9X7Y5	Q9X7Y5 streptomyce
32	8	2.1	626	16	Q9A2Z3	Q9A2Z3 caulobacter
33	8	2.1	644	16	Q92KK6	Q92KK6 rhizobium m
34	8	2.1	669	17	Q9ZPV17	Q9ZPV17 methanosarc
35	8	2.1	721	1	Q9V2S8	Q9V2S8 pyrobaculum
36	8	2.1	721	17	Q8ZW18	Q8ZW18 pyrobaculum
37	8	2.1	1179	17	Q59462	Q59462 pyrococcus
38	8	2.1	3063	5	Q95Z82	Q95Z82 leishmania
39	7	1.8	18	2	Q06514	Q06514 escherichia
40	7	1.8	19	10	Q9S8N3	Q9S8N3 triticum ae
41	7	1.8	68	16	Q8KBF2	Q8KBF2 chlorobium
42	7	1.8	77	10	Q8H8T3	Q8H8T3 oryza sativ
43	7	1.8	80	9	Q03961	Q03961 bacterioph
44	7	1.8	82	16	Q8YP38	Q8YP38 anabaena sp
45	7	1.8	91	2	Q91UN6	Q91UN6 rhizobium m

## ALIGNMENTS

### RESULT 1

Q9D695  
ID Q9D695 PRELIMINARY; PRT; 380 AA.  
AC Q9D695  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE 4631416M05Rik protein (Megsin).  
GN SERPINB7 OR 431416M05RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi, Muridae; Murinae; Mus.  
OX NCBI TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CS7BL/6J; TISSUE=Skin;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyrshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
RN [2]  
RP SEQUENCE FROM N.A.

```

RX MEDLINE=21368006; PubMed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangio proliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AK014524; BA329410.1; -.
DR EMBL; AF105328; AAL16768.1; -.
DR HSP; P05121; 1A7C.
DR MGD; MGI:2151053; Serpinb7.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF000079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCFB9CF4 CRC64;

Query Match 4.7%; Score 18; DB 11; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWVLVNAVYFKGW 181
Db 164 SSSAVWVLVNAVYFKGW 181
|||||

RESULT 2
Q920J5 PRELIMINARY; PRT; 380 AA.
AC Q920J5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Megsin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21368006; PubMed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangio proliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF105329; AAL16769.1; -.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF000079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 42821 MW; D8076CABEE2C2FBC CRC64;

Query Match 4.7%; Score 13; DB 11; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 SSSAVWVLVNAVYFKGW 181
Db 164 SSSAVWVLVNAVYFKGW 181
|||||

RESULT 3
Q9DHG4 PRELIMINARY; PRT; 334 AA.
AC Q9DHG4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 149R protein.

GN 149R.
OS Yaba-like disease virus (YLDV).
OC Viruses; dsRNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21176366; PubMed=11277691;
RA Lee H.J., Essani K., Smith G.L.;
RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus.";
RL Virology 281:170-192(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AJ293568; CAC21387.1; -.
DR HSP; P05120; 1B77.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF000079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 334 AA; 38190 MW; D7FD004191B3C19F CRC64;

Query Match 3.7%; Score 14; DB 12; Length 334;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 HKSYIEVTEGTEA 336
Db 275 HKSYIEVTEGTEA 288
|||||

RESULT 4
Q8VHQO PRELIMINARY; PRT; 385 AA.
AC Q8VHQO;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SPI3L2.
GN C76171 OR SPI3L2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6D2 F1/J;
RA Kaisekman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,
RA McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,
RA Ragoussis J., Bird P.I.;
RT "Comparison of Human Chromosome 6p25 with Murine Chromosome 13 Reveals
RT a Greatly Expanded Ovary-Serpin Gene Repertoire in the Mouse.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF425084; AAL65910.1; -.
DR HSP; P01009; IQUP.
DR MGD; MGI:2145481; C76171.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF000079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 385 AA; 43786 MW; C820A9A26ABEB485 CRC64;

Query Match 2.9%; Score 11; DB 11; Length 385;
Best Local Similarity 100.0%; Pred. No. 0.013;
```

```

GN 149R.
OS Yaba-like disease virus (YLDV).
OC Viruses; dsRNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21176366; PubMed=11277691;
RA Lee H.J., Essani K., Smith G.L.;
RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus.";
RL Virology 281:170-192(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AJ293568; CAC21387.1; -.
DR HSP; P05120; 1B77.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF000079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 334 AA; 38190 MW; D7FD004191B3C19F CRC64;

Query Match 3.7%; Score 14; DB 12; Length 334;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 HKSYIEVTEGTEA 336
Db 275 HKSYIEVTEGTEA 288
|||||

RESULT 4
Q8VHQO PRELIMINARY; PRT; 385 AA.
AC Q8VHQO;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SPI3L2.
GN C76171 OR SPI3L2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6D2 F1/J;
RA Kaisekman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,
RA McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,
RA Ragoussis J., Bird P.I.;
RT "Comparison of Human Chromosome 6p25 with Murine Chromosome 13 Reveals
RT a Greatly Expanded Ovary-Serpin Gene Repertoire in the Mouse.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF425084; AAL65910.1; -.
DR HSP; P01009; IQUP.
DR MGD; MGI:2145481; C76171.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF000079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 385 AA; 43786 MW; C820A9A26ABEB485 CRC64;

Query Match 2.9%; Score 11; DB 11; Length 385;
Best Local Similarity 100.0%; Pred. No. 0.013;
```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 LVNAVYFKGW 181  
 Db 165 LVNAVYFKGW 175

RESULT 5

Q922G2  
 ID Q922G2 PRELIMINARY; PRT; 388 AA.  
 AC Q922G2;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Squamous cell carcinoma antigen 2.  
 GN SERPINB4 OR SCCA2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1] \_  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99047536; PubMed=9828132;  
 RX Bartuski A.J., Kamachi Y., Schick C., Massa H., Trask E.J.,  
 RA Silverman G.A.;  
 RT "A murine ortholog of the human serpin SCCA2 maps to chromosome 1 and  
 RT inhibits chymotrypsin-like serine proteinases."  
 RL Genomics 54:297-306(1998).  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR EMBL; AF063937; AAC95432.1; -;  
 DR HSSP; P05120; 1BY7.  
 DR MGD; MGI:1277952; Serpinb4.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Protease inhibitor; Serine protease inhibitor; Serpin.  
 SQ SEQUENCE 388 AA; 44573 MW; D77A4FA1233EB9CD CRC64;

Query Match 2.98; Score 11; DB 11; Length 388;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 MVLNAVYFKG 179  
 Db 172 MVLNAVYFKG 182

RESULT 6

Q8WNW8  
 ID Q8WNW8 PRELIMINARY; PRT; 292 AA.  
 AC Q8WNW8;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Nexin-1 (Fragment).  
 GN PN-1.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 [1] \_  
 RN SEQUENCE FROM N.A.  
 RP Paradis V., Lussier J., Silversides D.W.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR EMBL; AY059415; AAL23838.2; -;  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Protease inhibitor; Serine protease inhibitor; Serpin.  
 FT NON\_TER 1 1

SQ SEQUENCE 292 AA; 32597 MW; 5F0ACFF1DF6DC502 CRC64;

Query Match 2.6%; Score 10; DB 6; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLVNAVYFKG 179  
 Db 77 VLVNAVYFKG 86

RESULT 7

Q9D6A7  
 ID Q9D6A7 PRELIMINARY; PRT; 359 AA.  
 AC Q9D6A7;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE 393042J058Rik protein.  
 GN SP111 OR 393042J058Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1] \_  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kado T., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weiss C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR EMBL; AK014448; BAB29358.1; -;  
 DR HSSP; P05120; 1BY7.  
 DR MGD; MGI:894669; Sp11.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Protease inhibitor; Serine protease inhibitor; Serpin.  
 SQ SEQUENCE 359 AA; 41091 MW; C1109BE74A4814BD CRC64;

Query Match 2.6%; Score 10; DB 11; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 ECGTEAAT 340  
 Db 307 ECGTEAAT 316

RESULT 8

Q9D1E7  
 ID Q9D1E7 PRELIMINARY; PRT; 386 AA.  
 AC Q9D1E7;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)

RESULT 11

```

Q8B396
ID Q8B396 PRELIMINARY; PRT; 387 AA.
AC Q8B396;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Squamous cell carcinoma antigen 2.
GN SERPINB4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SFRAIN-Car-S, and Car-P; TISSUE=TPA-treated skin;
RA Gariboldi M., Peissel B., Fabbri A., Saran A., Zaifaroni D.,
RA Spinola M., Falvella F.S., Pazzaglia S., Tanuma J.-I., Maurichi A.,
RA Bartoli C., Silverman G., Covelli V., Pilotti S., Hayashizaki Y.,
RA Okazaki Y., Dragani T.A.;
RT "The serpin Sca2 gene plays a functional role in genetic
RT susceptibility to skin tumorigenesis in mice and humans.";
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY144683; AAN62870.1; -.
DR EMBL; AY144684; AAN62871.1; -.
SQ SEQUENCE 387 AA; 44434 MW; B5B15F807FEF0688 CRC64;

Query Match 2.6%; Score 10; DB 11; Length 387;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNVAVYFKG 179
Db 172 VLNVAVYFKG 181
|||||

RESULT 12
Q8HZY1
ID Q8HZY1 PRELIMINARY; PRT; 397 AA.
AC Q8HZY1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Serine protease inhibitor-E2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC Bedard J., Brule S., Price C., Silversides D.W., Lussier J.;
RT "Serine protease inhibitor-E2 (SERPINE2) is differentially expressed
RT in granulosa cells of dominant follicle in cattle.";
RL Mol. Reprod. Dev. 0:0-0(2002).
DR EMBL; AF251153; AAN37922.1; -.
KW Protease.
SQ SEQUENCE 397 AA; 43877 MW; 90893C66D6A15443 CRC64;

Query Match 2.6%; Score 10; DB 6; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLNVAVYFKG 179
Db 182 VLNVAVYFKG 191
|||||

RESULT 13
Q8YF8
ID Q8YF8 PRELIMINARY; PRT; 246 AA.
AC Q8YF8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein AL10902.
GN AL10902.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2159285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Yakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003584; BAB72859.1; -.
DR InterPro; IPR001853; DSEA.
DR InterPro; IPR006663; ThioRedox_dom2.
DR Pfam; PF01323; DSEA; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 248 AA; 27809 MW; A5EFP1919A66E2DA CRC64;

Query Match 2.4%; Score 9; DB 16; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 DLSEIENKL 253
Db 236 DLSEIENKL 244
|||||

RESULT 14
Q9D7P9
ID Q9D7P9 PRELIMINARY; PRT; 423 AA.
AC Q9D7P9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 2300003F07RIK protein (Weakly similar to serpin B12).
GN 2300003F07RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King S., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Konteski S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;

```



```

RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AK09018; BAB26028.1; -.
DR EMBL; AK040697; BAC30672.1; -.
DR HSP; P05619; 1HLE.
DR MGD; MGI:1919119; 2300003P07Rik.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 423 AA; 47834 MW; 5A22BE2FE51B6120 CRC64;

Query Match 2.4%; Score 9; DB 11; Length 423;
Best Local Similarity 100.0%; Pred.No.1.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 VLVNAVYFK 178
DB 204 VLVNAVYFK 212
|||||

RESULT 15
Q92K34
ID Q92K34 PRELIMINARY; PRT; 454 AA.
AC Q92K34;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein R02208.
GN R02208 OR SMC01575.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capella D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9877-9882(2001).
DR EMBL; AL591790; CAC45787.1; -.
DR InterPro; IPR005490; ErfK Ybis YhmG.
DR Pfam; PF03734; ErfK Ybis YhmG; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 454 AA; 49052 MW; 6C25A2A60730763D CRC64;

Query Match 2.4%; Score 9; DB 16; Length 454;
Best Local Similarity 100.0%; Pred.No.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 LSLFALAL 41
DB 9 LSLFALAL 17
|||||

```

Search completed: December 12, 2003, 16:34:53  
Job time : 52 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:19:36 ; Search time 43 Seconds  
(without alignments)  
1402.700 Million cell updates/sec

Title: US-09-508-997A-2  
Perfect score: 1959  
Sequence: 1 MASLAANAARPCNLFREND.....FLVIRKDIILFSGKVSVP 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	1959	100.0	380	15 AAR48379	Human megakaryocyte
2	1959	100.0	380	20 AAY08254	Human megasin prote
3	1959	100.0	380	21 AAB24142	Human megasin prote
4	1959	100.0	380	22 AAB83075	Human megasin prote
5	1959	100.0	380	24 ABU56533	Lung cancer-associ
6	1488	76.0	380	22 AAG64286	Rat megasin protein
7	1477	75.4	380	20 AAY08255	Rat megasin protein
8	1477	75.4	380	21 AAB24150	Rat megasin protein
9	1477	75.4	380	22 AAB83076	Rat megasin protein

10	1458	74.4	368	20 AAY08256	Mouse megasin prote
11	1458	74.4	368	21 AAB24151	Mouse megasin prote
12	1458	74.4	368	22 AAB83077	Murine megasin prot
13	814	41.6	378	23 ABG76511	DNA encoding prote
14	796.5	40.7	379	21 AAY15355	Human prostate ser
15	771.5	39.4	395	22 AAM47207	Human NOV1 protein
16	769	39.3	394	23 ABB53280	Human polypeptide
17	754.5	38.5	425	23 ABG70630	Amino acid sequenc
18	754	38.5	390	24 ABB98808	SCCA1/SCCA2 fusion
19	752	38.4	390	13 AAR25276	SCC antigen. Synt
20	752	38.4	390	20 AAY32077	Hepatitis B virus
21	752	38.4	390	20 AAY25928	Human SCCA1 protei
22	752	38.4	390	22 AAY72654	Human squamous cel
23	752	38.4	390	24 ABB98810	SCCA1 protein. Ho
24	750	38.3	390	20 AAY32078	Hepatitis B virus
25	748	38.2	390	20 AAY25927	Human SCCA2 protei
26	748	38.2	390	22 AAY72655	Human squamous cel
27	748	38.2	390	24 ABU56633	Lung cancer-associ
28	748	38.2	390	24 ABB98811	SCCA2 protein. Ho
29	744	38.0	390	18 AAM15242	Psoriasisstatin type
30	744	38.0	390	24 ABB98870	SCCA1/SCCA2 fusion
31	740	37.8	390	18 AAM15241	Psoriasisstatin type
32	738.5	37.7	617	21 AAY15156	Human prostate ser
33	722.5	36.9	379	17 AAR94367	Human elastase inh
34	722.5	36.9	437	21 AAB43755	Human cancer assoc
35	720	36.8	390	24 ABB98671	SCCA1/SCCA2 fusion
36	717.5	36.6	379	16 AAR64159	Human elastase inh
37	716	36.5	400	23 ABB09506	Human hurpin/PI 13
38	715.5	36.5	391	20 AAY41029	Human lung tumor a
39	715.5	36.5	391	21 AAB11314	Human lung cancer-
40	715.5	36.5	391	22 AAY72651	Human headpin (for
41	715.5	36.5	391	23 ABB61866	Human lung cancer
42	715.5	36.5	391	23 ABB74946	Human lung tumor
43	715	36.5	400	20 AAY41030	Human lung tumor a
44	715	36.5	400	21 AAB11315	Human hurpin/PI 13
45	715	36.5	400	23 ABB09507	

# ALIGNMENTS

RESULT 1  
NAR48379  
ID AAR48379 standard; Protein; 380 AA.  
XX AC AAR48379;  
XX AC AAR48379;  
DT 25-MAR-2003 (updated)  
DT 16-AUG-1994 (first entry)  
XX DE Human megakaryocyte differentiation factor.  
XX DE Human megakaryocyte differentiation factor; MDF; thrombopoietin;  
XX KW haematopoietic stimulating factor; thrombocytopaenia; platelet;  
XX KW bone marrow transplantation; cancer chemotherapy.  
XX OS Homo sapiens.  
XX PN BP583884-A1.  
XX PD 23-FEB-1994.  
XX PF 19-JUL-1993; 93EP-0305654.  
XX PR 17-JUL-1992; 92JP-0212305.  
XX PR 04-MAR-1993; 93JP-0057339.  
XX PA (SUNR ) SUNTORY LTD.  
XX PA (TSUJ ) TSUJIMOTO M.  
XX PI Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;  
XX PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamaichi K;

DR WPI; 1994-058782/08.  
 DR N-PSDB; AAQ56670.  
 XX  
 PT New megakaryocyte differentiation factor - isolated from human  
 PT epidermoid carcinoma cells, used to treat conditions involving a  
 PT decrease in platelets  
 XX  
 XX Claim 7; Page 30-32; 47pp; English.  
 XX  
 CC Human MDF can be isolated from a culture of human epidermoid  
 CC carcinoma A431 cells in protein-free medium. The MDF stimulates  
 CC differentiation of megakaryocytes from myeloid cells in the presence  
 CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful  
 CC for treatment of diseases involving a decrease in platelet number  
 CC (esp. thrombocytopenia) such as occurs in bone marrow  
 CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kD by  
 CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 380 AA;  
 SQ  
 Query Match 100.0%; Score 1959; DB 15; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-169;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASLAAANAEEFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60  
 DB 1 MASLAAANAEEFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60  
 QY 61 NTASGYGNSSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120  
 DB 61 NTASGYGNSSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120  
 QY 121 LYDAKVERVDFTNLEDTNRNINRWVENETHGKIKNVI GEGGISSSAVMVLVNAVYFKG 180  
 DB 121 LYDAKVERVDFTNLEDTNRNINRWVENETHGKIKNVI GEGGISSSAVMVLVNAVYFKG 180  
 QY 181 WSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240  
 DB 181 WSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240  
 QY 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPQPKIEKNYEMKQYLALGLKIDF 300  
 DB 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPQPKIEKNYEMKQYLALGLKIDF 300  
 QY 301 DESKADLSIASGGRLYSRMMHKSIVTEEGTEATAATGSNIVEKQLPQSTLFRADHP 360  
 DB 301 DESKADLSIASGGRLYSRMMHKSIVTEEGTEATAATGSNIVEKQLPQSTLFRADHP 360  
 QY 361 FLFVIRKDDIILFSGKVSCP 380  
 DB 361 FLFVIRKDDIILFSGKVSCP 380  
 RESULT 2  
 AAY08254  
 ID AAY08254 standard; Protein; 380 AA.  
 XX  
 AC AAY08254;  
 XX  
 DT 14-JUL-1999 (first entry)  
 DE Human megsin protein.  
 XX  
 KW Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;  
 KW human; rat; murine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09915652-A1.  
 XX  
 PD 01-APR-1999.  
 XX

PF 22-SEP-1998; 98WO-JP04269.  
 XX  
 PR 22-SEP-1997; 97JP-0275302.  
 XX  
 PA (KUPO/) KUROKAWA K.  
 PA (MIYA/) MIYATA T.  
 XX  
 PI Miyata T;  
 XX  
 DR WPI; 1999-276983/23.  
 DR N-PSDB; AAX56712.  
 XX  
 PT Megsin protein expressed specifically in mesangial cells  
 XX  
 XX Claim 1; Page 62-64; 100pp; Japanese.  
 PS  
 CC This invention describes the isolation of novel megsin nucleic acid and  
 CC proteins from human, rat and mouse tissue. This protein is expressed  
 CC specifically in mesangial cells. The products of the invention are  
 CC useful for the treatment and diagnosis of diseases involving mesangial  
 CC cells, such as IgA nephropathy.  
 XX  
 XX Sequence 380 AA;  
 SQ  
 Query Match 100.0%; Score 1959; DB 20; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-169;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASLAAANAEEFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60  
 DB 1 MASLAAANAEEFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60  
 QY 61 NTASGYGNSSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120  
 DB 61 NTASGYGNSSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120  
 QY 121 LYDAKVERVDFTNLEDTNRNINRWVENETHGKIKNVI GEGGISSSAVMVLVNAVYFKG 180  
 DB 121 LYDAKVERVDFTNLEDTNRNINRWVENETHGKIKNVI GEGGISSSAVMVLVNAVYFKG 180  
 QY 181 WSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240  
 DB 181 WSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240  
 QY 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPQPKIEKNYEMKQYLALGLKIDF 300  
 DB 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPQPKIEKNYEMKQYLALGLKIDF 300  
 QY 301 DESKADLSIASGGRLYSRMMHKSIVTEEGTEATAATGSNIVEKQLPQSTLFRADHP 360  
 DB 301 DESKADLSIASGGRLYSRMMHKSIVTEEGTEATAATGSNIVEKQLPQSTLFRADHP 360  
 QY 361 FLFVIRKDDIILFSGKVSCP 380  
 DB 361 FLFVIRKDDIILFSGKVSCP 380  
 RESULT 3  
 AAB24142  
 ID AAB24142 standard; Protein; 380 AA.  
 XX  
 AC AAB24142;  
 XX  
 DT 30-JAN-2001 (first entry)  
 DE Human megsin protein sequence SEQ ID NO:2.  
 XX  
 KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;  
 KW IgA; immunoglobulin A; detection; renal function; renal disorder;  
 KW diagnosis; biological sample; blood; urine.  
 XX  
 OS Homo sapiens.  
 XX



QY 361 FLVIRKDDIILPSGKVSCP 380  
 DB 361 FLVIRKDDIILPSGKVSCP 380

## RESULT 5

ABU56533  
 ID ABU56533 standard; Protein; 380 AA.

XX AC ABU56533;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polypeptide #126.

XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US12476.

XX PR 18-APR-2001; 2001US-284770P.

XX PR 10-MAY-2001; 2001US-290492P.

XX PR 09-NOV-2001; 2001US-339245P.

XX PR 13-NOV-2001; 2001US-350866P.

XX PR 29-NOV-2001; 2001US-334370P.

XX PR 12-APR-2002; 2002US-372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;

XX DR WPI; 2003-093161/08.

XX DR N-PSDB; ABX76260.

XX PT Detecting a lung cancer-associated transcript in a cell from a patient  
 PT for treating lung cancer, by contacting a biological sample from the  
 PT patient with a polynucleotide that exhibits increased or decreased  
 XX expression in lung cancer

XX PS Claim 27; Page 288; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated  
 CC transcript in a cell from a patient, comprising contacting a biological  
 CC sample from the patient with a polynucleotide that selectively hybridises  
 CC to a sequence that is at least 80 % identical to a gene that exhibits  
 CC increased or decreased expression in lung cancer samples. Lung  
 CC cancer-associated polynucleotides and polypeptides are used for  
 CC identifying a compound that modulates a lung cancer-associated  
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated  
 CC cell to treat lung cancer in a patient and for treating a mammal having  
 CC lung cancer by administering a modulatory compound identified. The  
 CC methods are useful for treating lung cancer, such as small cell lung  
 CC cancer, non-small cell lung cancer or other benign or precancerous  
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive  
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial  
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides  
 CC and polypeptides are useful for diagnostic purposes and as targets for  
 CC screening for therapeutic compounds that modulate lung cancer, such as  
 CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated  
 CC polypeptides of the invention.

XX SQ Sequence 380 AA;

Query Match 100.0%; Score 1959; DB 24; Length 380;

Best Local Similarity 100.0%; Pred. No. 3.4e-169;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASIAAANAEECFNLPRMDNQGNGNVFSSLSLFAALALVRLGAQDLSLQIDKLHV 60  
 DB 1 MASIAAANAEECFNLPRMDNQGNGNVFSSLSLFAALALVRLGAQDLSLQIDKLHV 60  
 QY 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120  
 DB 61 NTASGYGNSNSQSGLSQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120  
 QY 121 LYDAKVERVDFTNHLEDTNRNINKWVENETHGKTKNVIGEGGSISSAAMVLVNAVYFKG 180  
 DB 121 LYDAKVERVDFTNHLEDTNRNINKWVENETHGKTKNVIGEGGSISSAAMVLVNAVYFKG 180  
 QY 181 WQSAFTKSTTINCHFPKPSGSKAVAMVHQERKFNLSVIEDPSMKILEYNGINGINNYVL 240  
 DB 181 WQSAFTKSTTINCHFPKPSGSKAVAMVHQERKFNLSVIEDPSMKILEYNGINGINNYVL 240  
 QY 241 LPENDLSBIENKLTQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRALGLKDF 300  
 DB 241 LPENDLSBIENKLTQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRALGLKDF 300  
 QY 301 DESKADLSGIASGGRLYISMVHKSYIEVTEGTEATAATGTSNIVEKOLPQSTLFRADHP 360  
 DB 301 DESKADLSGIASGGRLYISMVHKSYIEVTEGTEATAATGTSNIVEKOLPQSTLFRADHP 360  
 QY 361 FLVIRKDDIILPSGKVSCP 380  
 DB 361 FLVIRKDDIILPSGKVSCP 380

## RESULT 6

AAG64286

ID AAG64286 standard; Protein; 380 AA.

XX AC AAG64286;

XX DT 21-SEP-2001 (first entry)

XX DE Rat meginin protein.

XX KW Rat; meginin; renal mesangial cell; mesangium proliferative nephritis.

XX OS Rattus norvegicus.

XX PN WO200148019-A1.

XX PD 05-JUL-2001.

XX PF 26-DEC-2000; 2000WO-JP09251.

XX PR 28-DEC-1999; 99JP-0373677.

XX PA (KURO/) KUROKAWA K.

XX PA (MIYA/) MIYATA T.

XX PI Miyata T;

XX DR WPI; 2001-425651/45.

XX DR N-PSDB; AA#48181.

XX PT New antibody recognizing a partial sequence of rat meginin protein for  
 PT diagnosis of mesangium proliferative nephritis

XX PS Disclosure; Page 54-56; 63pp; Japanese.

XX CC The present invention relates to a novel antibody which recognises a  
 CC peptide consisting of residues 341-354 of rat meginin protein. The present  
 CC sequence is the protein sequence for rat meginin, which was used in the  
 CC present invention. Meginin is highly expressed in renal mesangial cells  
 CC and its level is elevated in mesangium proliferative nephritis. Assay of  
 CC the serum or urine level using the antibody is therefore indicative of

CC this type of disorder.

XX Sequence 380 AA;

Query Match 76.0%; Score 1488; DB 22; Length 380;

Best Local Similarity 73.7%; Pred. No. 2e-126;

Matches 280; Conservative 57; Mismatches 43; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNQNGNVFSSLSFAALALVRLGAQDDLSQIDKLHV 60  
 DB 1 MASLAANAEEFCNLFREMDNQNGNVFSSLSFAALALVRLGAQDDLSQIDKLHV 60  
 QY 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSVNGLFAEKVYGFHKDYIECAEK 120  
 DB 61 ISPRQNGNSNSQSGLOQLKRVFSDINASHKDYDLSVNGLFAEKVYGFHKDYIECAEK 120  
 QY 121 LYDAKVERVDFTHLEDFRININKWVENETHGKIKNVIKGGISSAVMVLVNAVYFKGK 180  
 DB 121 LYNKAKVERVDFTHLEDFRININKWVENETHGKIKNVIKGGISSAVMVLVNAVYFKGK 180  
 QY 181 WQSAPTKSETINCHFKSPKSGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240  
 DB 181 WKSAPTKSETINCHFKSPKSGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240  
 QY 241 LPENDLSEIENKLTQNLMEWTPRRMTSKYVEFFQPKIEKNYEMKQYLALGLKDIF 300  
 DB 241 LPEDDLSEIESKLSFQNLMDWTSRKMKSYVNVFLPQPKIEKNYEMKQYLALGLKDIF 300  
 QY 301 DESKADLSGIASGGRLYVSKLMHKSLEIVSEGTETAATAATGSNIVEKQLPOSTLFRADHP 360  
 DB 301 VESRADLSGIASGGRLYVSKLMHKSLEIVSEGTETAATAATGSNIVEKQLPOSTLFRADHP 360  
 QY 361 FLFVIRKDDIILFSGKVSFCP 380  
 DB 361 FLFVIRKNGIILFTGKVSFCP 380

RESULT 7

AAV08255

ID AAY08255 standard; Protein; 380 AA.

AC AAY08255;

XX 14-JUL-1999 (first entry)

DT Rat mesgin protein.

DE Mesgin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;

KW human; rat; murine.

XX Rattus rattus.

OS WO9915652-A1.

XX 01-APR-1999.

XX 22-SEP-1998; 98WO-JP04269.

XX 22-SEP-1997; 97JP-0275302.

XX (KURO//) KUROKAWA K.

XX (MIYA//) MIYATA T.

XX Miyata T;

XX WPI; 1999-276983/23.

XX N-PSDB; AAX56712.

XX Mesgin protein expressed specifically in mesangial cells

PT Claim 1; Page 69-72; 100pp; Japanese.

XX This invention describes the isolation of novel mesgin nucleic acid and

CC

CC proteins from human, rat and mouse tissue. This protein is expressed  
 CC specifically in mesangial cells. The products of the invention are  
 CC useful for the treatment and diagnosis of diseases involving mesangial  
 CC cells, such as IgA nephropathy.

XX Sequence 380 AA;

Query Match 75.4%; Score 1477; DB 20; Length 380;

Best Local Similarity 73.4%; Pred. No. 2e-125;

Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNQNGNVFSSLSFAALALVRLGAQDDLSQIDKLHV 60  
 DB 1 MASLAANAEEFCNLFREMDNQNGNVFSSLSFAALALVRLGAQDDLSQIDKLHV 60  
 QY 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSVNGLFAEKVYGFHKDYIECAEK 120  
 DB 61 ISPRQNGNSNSQSGLOQLKRVFSDINASHKDYDLSVNGLFAEKVYGFHKDYIECAEK 120  
 QY 121 LYDAKVERVDFTHLEDFRININKWVENETHGKIKNVIKGGISSAVMVLVNAVYFKGK 180  
 DB 121 LYNKAKVERVDFTHLEDFRININKWVENETHGKIKNVIKGGISSAVMVLVNAVYFKGK 180  
 QY 181 WQSAPTKSETINCHFKSPKSGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240  
 DB 181 WKSAPTKSETINCHFKSPKSGKAVAMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240  
 QY 241 LPENDLSEIENKLTQNLMEWTPRRMTSKYVEFFQPKIEKNYEMKQYLALGLKDIF 300  
 DB 241 LPEDDLSEIESKLSFQNLMDWTSRKMKSYVNVFLPQPKIEKNYEMKQYLALGLKDIF 300  
 QY 301 DESKADLSGIASGGRLYVSKLMHKSLEIVSEGTETAATAATGSNIVEKQLPOSTLFRADHP 360  
 DB 301 VESRADLSGIASGGRLYVSKLMHKSLEIVSEGTETAATAATGSNIVEKQLPOSTLFRADHP 360  
 QY 361 FLFVIRKDDIILFSGKVSFCP 380  
 DB 361 FLFVIRKNGIILFTGKVSFCP 380

RESULT 8

AAV24150

ID AAB24150 standard; Protein; 380 AA.

AC AAB24150;

XX 30-JAN-2001 (first entry)

DT Rat mesgin protein sequence SEQ ID NO:19.

DE Mesgin; mesangium-predominant gene; serpin regulated; nephropathy;

KW IgA; immunoglobulin A; detection; renal function; renal disorder;

XX diagnosis; biological sample; blood; urine.

XX Rattus norvegicus.

XX Key Location/Qualifiers

XX Misc-difference 51 /note= "unspecified"

XX Misc-difference 94 /note= "unspecified"

XX WO200057189-A1.

XX 28-SEP-2000.

XX 17-MAR-2000; 2000WO-JP01646.

XX 19-MAR-1999; 99JP-0075305.

XX 28-OCT-1999; 99JP-0306623.

XX (KURO//) KUROKAWA K.

XX (FUSO ) FUSO PHARM IND LTD.

```
PA (MIYA/) MIYATA T.
XX PI Miyata T;
XX PD WPI; 2000-611642/58.
XX DR N-PSDB; AAC55238.
XX PT Evaluating renal function comprises assaying megin protein in
XX PT biological sample -
XX PS Example 2; Page 81-84; 93pp; Japanese.
XX CC The present invention describes a method for evaluating renal function.
XX CC The method comprises assaying megin protein in biological sample. Also
XX CC described are: (1) use of an anti-megsin protein antibody for diagnosing
XX CC renal function; and (2) a kit for detecting megin protein comprising:
XX CC (a) anti-megsin protein antibody attached to solid magnetic particles;
XX CC (b) direct or indirect fixing for the antibody to the particles; and
XX CC (c) a magnet. The process is useful for evaluating renal function and
XX CC diagnosing renal disorders by assaying megin protein in biological
XX CC samples (preferably urine or blood). The process is reproducible and
XX CC gives accurate results. The present sequence represents the rat megin
XX CC protein, which is given in the exemplification of the present invention.
XX SQ Sequence 380 AA;

Query Match 75.4%; Score 1477; DB 21; Length 380;
Best Local Similarity 73.4%; Pred. No. 2e-125;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAQDDSLSQIDKLHY 60
Db 1 MASLAANAEEFGDLFREMDSSQNGNVFFSSLSLFTALSIRLGARGDCXRQDKALHF 60

QY 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
Db 61 ISPSRQGNSSNQLGLQYQKRVLADINSHKHDXKLSIANGVFAEKVDFPHKSYMECAEN 120

QY 121 LYDAKVERVDFTNHLDETRRNINKWVENETHGKIKNVIKNGVIGSGISSAVVNLVNAVYFGK 180
Db 121 LYNAKVERVDFTNDIQETRFKINKNIENETHGKIKVLDGSSLSLSSAVVNLVNAVYFGK 180

QY 181 WSAFTKSTIRCHPKSCGKAVAMHQRKFNLSVIEDPSMKILEIRYNGINMYL 240
Db 181 WSAFTKSTIRCHPKSCGKAVAMHQRKFNLSVIEDPSMKILEIRYNGINMYL 240

QY 241 LPENDLSEIKLTONLMEWTNPRMTSKYVEVFFPQKIEKNYEMKQYLGKIDIF 300
Db 241 LPENDLSEIKLTONLMEWTNPRMTSKYVEVFFPQKIEKNYEMKQYLGKIDIF 300

QY 301 DESKADLSGIASGGLYISRMHKSIVETBEGTEATAATGSNIVEKQIPQSTLFRADHP 360
Db 301 VESRADLSGIASGGLYISRMHKSIVETBEGTEATAATGSNIVEKQIPQSTLFRADHP 360

QY 361 FLFVIRKDDIILFSGKVSCEP 380
Db 361 FLFVIRKNGIILFTGKVSCEP 380

RESULT 9
AAB83076
ID AAB83076 standard; Protein; 380 AA.
AC AAB83076;
XX 10-JUL-2001 (first entry)
XX Rat megin protein.
XX Rat; megin; mesangial cell proliferative nephritis; nephrotropic;
XX transgenic mouse; glomerular disease; animal model; drug screening.
XX Os Rattus norvegicus.
```

```
XX WO200124628-A1.
XX 12-APR-2001.
XX 06-OCT-2000; 2000WO-JP06988.
XX 06-OCT-1999; 99JP-0285736.
XX (KURO/) KUROKAWA K.
XX (MIYA/) MIYATA T.
XX Miyata T;
XX WPI; 2001-300136/31.
XX N-PSDB; AAF82439.
XX Mouse model for mesangial cell proliferative nephritis for development
XX PT and screening of new treatments -
XX PS Disclosure; Page 48-50; 62pp; Japanese.
XX CC The present sequence is rat megin. The human megin coding
XX CC sequence may be introduced into a mouse to produce an animal model of
XX CC mesangial cell proliferative nephritis. The symptoms include
XX CC enlargement of the mesangial base region, sedimentation of an immune
XX CC complex and an increase in mesangial cells. The animal model is useful
XX CC for analysing the pathology of chronic glomerular diseases and for
XX CC screening compositions for prevention and treatment of the diseases.
XX CC Highly uniform models can be made easily and in large numbers using
XX CC this method.
XX SQ Sequence 380 AA;

Query Match 75.4%; Score 1477; DB 22; Length 380;
Best Local Similarity 73.4%; Pred. No. 2e-125;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAQDDSLSQIDKLHY 60
Db 1 MASLAANAEEFGDLFREMDSSQNGNVFFSSLSLFTALSIRLGARGDCXRQDKALHF 60

QY 61 NTASGYGNSNSQSGLOQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEK 120
Db 61 ISPSRQGNSSNQLGLQYQKRVLADINSHKHDXKLSIANGVFAEKVDFPHKSYMECAEN 120

QY 121 LYDAKVERVDFTNHLDETRRNINKWVENETHGKIKNVIKNGVIGSGISSAVVNLVNAVYFGK 180
Db 121 LYNAKVERVDFTNDIQETRFKINKNIENETHGKIKVLDGSSLSLSSAVVNLVNAVYFGK 180

QY 181 WSAFTKSTIRCHPKSCGKAVAMHQRKFNLSVIEDPSMKILEIRYNGINMYL 240
Db 181 WSAFTKSTIRCHPKSCGKAVAMHQRKFNLSVIEDPSMKILEIRYNGINMYL 240

QY 241 LPENDLSEIKLTONLMEWTNPRMTSKYVEVFFPQKIEKNYEMKQYLGKIDIF 300
Db 241 LPENDLSEIKLTONLMEWTNPRMTSKYVEVFFPQKIEKNYEMKQYLGKIDIF 300

QY 301 DESKADLSGIASGGLYISRMHKSIVETBEGTEATAATGSNIVEKQIPQSTLFRADHP 360
Db 301 VESRADLSGIASGGLYISRMHKSIVETBEGTEATAATGSNIVEKQIPQSTLFRADHP 360

QY 361 FLFVIRKDDIILFSGKVSCEP 380
Db 361 FLFVIRKNGIILFTGKVSCEP 380

RESULT 10
AAY08256
ID AAY08256 standard; Protein; 368 AA.
XX AC AAY08256;
XX
```

DT 14-JUL-1999 (first entry)  
 XX Mouse megin protein.  
 DE  
 XX Megin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;  
 KW human; rat; murine.  
 XX  
 OS Mus musculus.  
 XX  
 XX Key Location/Qualifiers  
 FT Protein 1..368  
 FT /note= "partial sequence"  
 XX  
 XX WO9515652-A1.  
 PN  
 XX 01-APR-1999.  
 PD  
 XX 22-SEP-1998; 98WO-JP04269.  
 PF  
 XX 22-SEP-1997; 97JP-0275302.  
 PR  
 XX (KURO/) KUROKAWA K.  
 PA (MIYA/) MIYATA T.  
 XX  
 XX Miyata T;  
 PI  
 XX WPI; 1999-276983/23.  
 DR  
 XX N-PSDB; AAS56714.  
 PA  
 XX Megin protein expressed specifically in mesangial cells  
 FT  
 XX Claim 1; Page 76-79; 100pp; Japanese.  
 PS  
 XX This invention describes the isolation of novel megin nucleic acid and  
 CC proteins from human, rat and mouse tissue. This protein is expressed  
 CC specifically in mesangial cells. The products of the invention are  
 CC useful for the treatment and diagnosis of diseases involving mesangial  
 CC cells, such as IgA nephropathy.  
 XX  
 SQ Sequence 368 AA;  
 XX  
 Query Match 74.4%; Score 1458; DB 20; Length 368;  
 Best Local Similarity 73.4%; Pred. No. 1e-123;  
 Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;  
 QY 13 FNLFREMDNCGNVPFFSLFPAALALVRLGAQDDSLQDKLLHNTASGYGNSNS 72  
 Db 1 FDLFREMDSSQGNVPFFSLFPAALALVRLGAQDDSLQDKLLHNTASGYGNSNS 60  
 QY 73 QGLOSQKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAKLYDAKVERVDFT 132  
 Db 61 QGLOVQLKRVLADINSSHKDYELSIATGVFAEKVYDFHKYIECAENLNKVERVDFT 120  
 QY 133 NHELTRNINKWVENETGKIKWVIGEGGISSAWVLVNAVYFKGWSAFTKSETIN 192  
 Db 121 NDVQDTRFKINKWVENETGKIKWVIGEGGISSAWVLVNAVYFKGWSAFTKSETIN 180  
 QY 193 CHFSPKSGKAVAMHQRKFNLSVIEDPSMKILRLYNGINGINMYVLLPNDLSEIENK 252  
 Db 181 CFRSPKSGKAVAMHQRKFNLSVIEDPSMKILRLYNGINGINMYVLLPNDLSEIENK 240  
 QY 253 LTFQNLMEWTNPRMTSKVEVFPQPKIEKNYEMQVLRALGKIDFDESKADLSGTAS 312  
 Db 241 LTFQNLMEWTNPRMTSKVEVFPQPKIEKNYEMQVLRALGKIDFDESKADLSGTAS 300  
 QY 313 GGRLYSRMKHSYEVTEGTEATAATGSNIVEKOLPQSTLFRADHPFLFVIRKDDIIL 372  
 Db 301 GGRLYSRMKHSYEVTEGTEATAATGSNIVEKOLPQSTLFRADHPFLFVIRKDDIIL 360  
 QY 373 FSGKVSFC 380  
 Db 361 FTGKVSFC 368

RESULT 11  
 AAB24151  
 ID AAB24151 standard; Protein; 368 AA.  
 XX  
 AC AAB24151;  
 XX  
 DT 30-JAN-2001 (first entry)  
 XX  
 XX Mouse megin protein sequence SEQ ID NO:21.  
 DE  
 XX Megin; mesangium-predominant gene; serpin regulated; nephropathy;  
 KW IgA; immunoglobulin A; detection; renal function; renal disorder;  
 KW diagnosis; biological sample; blood; urine.  
 XX  
 OS Mus musculus.  
 XX  
 XX WO2000057189-A1.  
 PN  
 XX 28-SEP-2000.  
 PD  
 XX 17-MAR-2000; 2000WO-JP01646.  
 PF  
 XX 19-MAR-1999; 99JP-0075305.  
 PR  
 XX 28-OCT-1999; 99JP-0306623.  
 PA (KURO/) KUROKAWA K.  
 PA (FUSO) FUSO PHARM IND LTD.  
 PA (MIYA/) MIYATA T.  
 XX  
 XX Miyata T;  
 PI  
 XX WPI; 2000-611642/58.  
 DR  
 XX N-PSDB; AACS5239.  
 PA  
 XX Evaluating renal function comprises assaying megin protein in  
 FT biological sample -  
 FT  
 XX Disclosure; Page 89-91; 93pp; Japanese.  
 PS  
 XX The present invention describes a method for evaluating renal function.  
 CC The method comprises assaying megin protein in biological sample. Also  
 CC described are: (1) use of an anti-megin protein antibody for diagnosing  
 CC renal function; and (2) a kit for detecting megin protein comprising:  
 CC (a) anti-megin protein antibody attached to solid magnetic particles;  
 CC (b) direct or indirect fixing for the antibody to the particles; and  
 CC (c) a magnet. The process is useful for evaluating renal function and  
 CC diagnosing renal disorders by assaying megin protein in biological  
 CC samples (preferably urine or blood). The process is reproducible and  
 CC gives accurate results. The present sequence represents the mouse megin  
 CC protein, which is given in the exemplification of the present invention.  
 XX  
 SQ Sequence 368 AA;  
 XX  
 Query Match 74.4%; Score 1458; DB 21; Length 368;  
 Best Local Similarity 73.4%; Pred. No. 1e-123;  
 Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;  
 QY 13 FNLFREMDNCGNVPFFSLFPAALALVRLGAQDDSLQDKLLHNTASGYGNSNS 72  
 Db 1 FDLFREMDSSQGNVPFFSLFPAALALVRLGAQDDSLQDKLLHNTASGYGNSNS 60  
 QY 73 QGLOSQKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAKLYDAKVERVDFT 132  
 Db 61 QGLOVQLKRVLADINSSHKDYELSIATGVFAEKVYDFHKYIECAENLNKVERVDFT 120  
 QY 133 NHELTRNINKWVENETGKIKWVIGEGGISSAWVLVNAVYFKGWSAFTKSETIN 192  
 Db 121 NDVQDTRFKINKWVENETGKIKWVIGEGGISSAWVLVNAVYFKGWSAFTKSETIN 180  
 QY 193 CHFSPKSGKAVAMHQRKFNLSVIEDPSMKILRLYNGINGINMYVLLPNDLSEIENK 252  
 Db 181 CFRSPKSGKAVAMHQRKFNLSVIEDPSMKILRLYNGINGINMYVLLPNDLSEIENK 240





PT New protein modification and maintenance molecules useful for treating  
 PT or preventing gastrointestinal, cardiovascular,  
 PT autoimmune/inflammatory, cell proliferative, developmental,  
 PT neurological and reproductive disorders -  
 XX  
 PS Claim 1; Page 154-155; 172pp; English.

XX  
 CC The invention describes an isolated human polypeptide (I), a naturally  
 CC occurring amino acid sequence at least 90 % identical to the protein, or  
 CC a biologically active fragment or an immunogenic fragment of the protein.  
 CC The protein modification and maintenance molecules are useful in the  
 CC diagnosis, treatment, and prevention of gastrointestinal (e.g. dysphagia,  
 CC esophageal spasm, gastritis, anorexia or nausea), cardiovascular (e.g.  
 CC atherosclerosis, hypertension, vasculitis, aneurysms, or ischemic heart  
 CC disease), autoimmune/inflammatory (e.g. acquired immunodeficiency  
 CC syndrome (AIDS), allergies, ankylosing spondylitis, anaemia or  
 CC amyloidosis), cell proliferative (e.g. cancers, arteriosclerotic,  
 CC bursitis, or cirrhosis), developmental (e.g. renal tubular acidosis,  
 CC anaemia, bone resorption, or epilepsy), epithelial (e.g. allergic contact  
 CC dermatitis, keratosis pilaris, insect bites, keloid, dermatofibroma or  
 CC eczema), neurological (e.g. stroke, cerebral neoplasms, Alzheimer's  
 CC disease), Huntington's disease or dementia), and reproductive disorders  
 CC (e.g. infertility, endometriosis, gynecomastia or ectopic pregnancy).  
 CC These may also be used in assessing the effects of exogenous compounds on  
 CC the expression of nucleic acid and amino acid sequences of protein  
 CC modification and maintenance molecules. Polynucleotides are useful in  
 CC somatic and germline gene therapy. This is the amino acid sequence of a  
 CC protein modification and maintenance molecule described in the invention.

XX Sequence 378 AA;

Query Match 41.6%; Score 814; DB 23; Length 378;  
 Best Local Similarity 42.6%; Pred. No. 3.4e-65;  
 Matches 164; Conservative 80; Mismatches 129; Indels 12; Gaps 4;

QY 1 MASLAANAFCFNLPREDDNQNGNVFFSSLSLFAALVRLGAQDDSLSDIKLHV 60  
 DB 1 MGSLSANVEFLDVFAELNSNNIGDNIFSSLSLYALSWLLGARGETAQLKLVHF 60  
 QY 61 NTASGYGNSNSQSGQLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKD 120  
 DB 61 SEVLHFSHTVDS-----LKPGFKDPSKPSDNCSTLSIANRLYTKTMAPHQQLSCSEK 113  
 QY 121 LYDAKVERVDFTNHEIDTRRNKXWVENETHGKIKNVIGEGISSAWMLVNAVYKCK 180  
 DB 114 WQARLQVDFEQASTETKTNANVENKNGKVALFGKSTIDPSSWMLVNAVYFKGQ 173  
 QY 181 WQATKGETINCHFKSPKCGKAVAMHQRXFNLSVIEDPFSKILELY-NGGINMYV 239  
 DB 174 WQKQVRETVKQFQPSLSEGNVTVMYQIGTFKLFAPKVPQVLELPYNNKLSMI 233  
 QY 240 LLPEN--ELSEIKKLTQNLKMTNPRMTSKYVFPFPQFKYKYMKQVLPALGLK 297  
 DB 234 LLPVGIAMKQLEKQNSGTFHETSSNNMEREVHLFPFKLEIKYELNSLLPLGVT 293  
 QY 298 DIPDSKADLSAGSGRLYISRMHKSYLEVTEETATATAGTSNVEKQLPQSTLERA 357  
 DB 294 DLFNQVKADLSGSPKGLYLSKAHKSYLDVSEETEAATGDSIAVKSPLMRAQFKA 353  
 QY 358 DHPFLFVIR--KDDILFGKVSQCP 380  
 DB 354 NHPFLFFIRHTHTNITLFGKLSAP 378

RESULT 14

AA15155  
 ID AA15155 standard; Protein; 379 AA.

XX  
 AC AA15155;

XX  
 DT 07-FEB-2000 (first entry)

XX  
 DE Human prostate serine protease inhibitor (PROSTAPIN).

XX  
 KW PROSTAPIN; serine protease inhibitor; serpin family; prostate; human;  
 KW prostate-specific tumour suppressor; apoptosis-inducer; diagnosis;  
 KW apoptosis-modulator; prognosis; treatment; prostate cancer;  
 KW metastatic prostate cancer.

XX Homo sapiens.

XX Key Location/Qualifiers  
 XX Inhibitory-site 330..347

FT /label= Serine protease inhibitory domain  
 FT /note= "Also known as reactive-site loop (RSL) and  
 FT contains a hinge region and a variable region. RSL  
 FT binds to protease active site during inhibition"  
 FT 330..338

FT Region /label= RSL hinge region

FT /note= "confers stability to the serpin-protease  
 FT complex and is highly conserved among the serpins"  
 FT 344..345

FT Cleavage-site  
 FT /note= "included in the RSL variable region. PROSTAPIN  
 FT undergoes cleavage at this site during inhibition"

XX WO9598560-A2.

XX 18-NOV-1999.

XX 31-WAR-1999; 99WO-US07123.

XX 31-WAR-1998; 98US-0080167.

XX 15-MAY-1998; 98US-0085720.

XX (UROC-) UROGENESYS INC.

XX (AFAR/) AFAR D E.

XX (HUBE/) HUBERT R S.

XX (LEON/) LEONG K.

XX (RAIT/) RAITANO A B.

XX (SAFF/) SAFFRAN D C.

XX Afar DE, Hubert RS, Leong K, Raitano AB, Safran DC;

XX WFI; 2000-039068/03.

XX N-PSDB; AA229049.

XX New isolated prostapin gene, used to develop products for, e.g.  
 XX treating and determining susceptibility to prostate cancer -

XX Claim 1; Fig 1; 63pp; English.

XX The present sequence is that of a PROSTAPIN protein which is a member of  
 XX serine protease inhibitor (serpin) family and expressed in normal  
 XX prostate and locally confined prostate cancer. The sequence is derived  
 XX from overlapping sequences of human cDNA clones 11P2A6 and 103. PROSTAPIN  
 XX contains a reactive-site loop characteristic of serpin family in  
 XX the carboxy-terminal region. The protein may function as a prostate-  
 XX specific tumour suppressor, apoptosis-inducer or apoptosis-modulator.  
 XX The sequence is useful in diagnosing, prognosing, monitoring and  
 XX treating prostate cancer, particularly advanced stage and metastatic  
 XX prostate cancer.

XX Sequence 379 AA;

Query Match 40.7%; Score 796.5; DB 21; Length 379;  
 Best Local Similarity 42.6%; Pred. No. 1.3e-63;

Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 1 MASLAANAFCFNLPREDDNQNGNVFFSSLSLFAALVRLGAQDDSLSDIKLHV 60

DB 1 MGSLSANVEFLDVFAELNSNNIGDNIFSSLSLYALSWLLGARGETAQLKLVHF 60

QY 61 -----NTASGYGNSN-SQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKD 113

DB 61 SHVDSLKPQKSPKSCQAGRIHSEGFVSQINQPSNCTLSIANRLYGTWAFHQQ 120

QY	114	YIECEKLYDAKVERVDFNFHLEDTNRNINKWENETHCKIKUVIGEGISSAVMYLVN	173
Db	121	YLSCEKWKYARLQTVDFEQSTETRMKNKAWYKTKGKVANLFGKSTIDPSSVMYLVN	180
QY	174	AVYFGKQWGAFTKSETINCHFPKPGCSGK--AVAMQERKFNLSVIEDFSMKILELY	231
Db	181	TIYFKGQNKPF-----QGNNVTVMYQIGTKFLAFVKPEQMVLELPY	225
QY	232	-NGGINVYVILPEN--DLSEINKLITCNLMETWTPRMTSKYVEVFFQFKIKKNYEMK	288
Db	226	VNNKLSMILLPVGIANLQIEKQKNSCTPHEWTSNNMREVEVHLFPFKLRTIKYELN	285
QY	289	QYLRALGKLDIPESKADLSFGAGRLYIYSRMKHSYIEVTRGTEATAATGNSIVKEQ	348
Db	286	SLLKPLGVTDLFQVQKADLSGNSPTKGLYLSKAIKHSILDVSEGTETAAATGDSIAVKS	345
QY	349	LQSTYLPADHPFLFVIR--XDDIILFSGKVSCP	380
Db	346	LPWRAQGFANHPFLFTRHTTNTILFCGLASP	379
RESULT 15			
AAW47207			
XX	ID	AAW47207 standard; Protein; 395 AA.	
XX	AC	AAW47207;	
DT	12-FEB-2002	(first entry)	
XX	DE	Human NOV1 protein.	
XX	DE	Human; NOV1; metabolic disorder; neurodegenerative disorder;	
KW	KW	immune disorder; haematopoietic disorder; developmental disease; cancer;	
KW	KW	retinal disease; feeding disorder; vaccine; infection; gene therapy;	
KW	KW	neurological disorder; psychotic disorder; G-protein coupled receptor;	
KW	KW	cytostatic; antidiabetic; viruicide; neuroprotective; nootropic;	
KW	KW	analgesic; antidepressant; antimigraine; anticonvulsant; neuroleptic;	
KW	KW	antiasthmatic; anti-allergic; antiinflammatory; anorectic; antiarthritic;	
KW	KW	antipeptic; antiatherosclerotic; antibacterial; fungicide;	
KW	KW	osteopathic; procoagulate; antitumor; hypertensive; hypotensive;	
KW	KW	antinfertility; vulnary; nephrotropic; antileptic; leupin;	
XX	XX	chromosome 18.	
OS	XX	Homo sapiens.	
PN	XX	WO200174851-A2.	
PD	PD	11-OCT-2001.	
PF	PF	30-MAR-2001; 2001WO-US10039.	
XX	XX	30-MAR-2000; 2000US-193205P.	
PR	PR	30-APR-2000; 2000US-193339P.	
PR	PR	05-APR-2000; 2000US-195343P.	
PR	PR	06-APR-2000; 2000US-195005P.	
PR	PR	06-APR-2000; 2000US-195082P.	
PR	PR	10-APR-2000; 2000US-195792P.	
PR	PR	11-APR-2000; 2000US-196556P.	
PR	PR	13-APR-2000; 2000US-197081P.	
PR	PR	14-APR-2000; 2000US-197087P.	
PR	PR	14-APR-2000; 2000US-197525P.	
PR	PR	29-MAR-2001; 2001US-0823187.	
PA	PA	(CUBA-) CURAGEN CORP.	
XX	XX		
PI	PI	Majumder K, Spaderna SK, Taupier RJ, Padigar M, Burgess CE;	
PI	PI	Shinkets RA, Szytek KA, Liu X, Patturajan M, Gusev VV;	
DR	DR	WPI: 2001-626379/72.	
DR	DR	N-PSDB; ABA01980.	
XX	XX		
PT	PT	New G protein-coupled receptor related polypeptides and polynucleotides	
PT	PT	for diagnosis, prevention and treatment of metabolic,	

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:16:51 ; Search time 21 Seconds  
(without alignments)  
765.625 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 1959

Sequence: 1 MASLAANAEEFCNLFREMD.....FLPVIRKDDIILFGKVSQP 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 segs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/aaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/aaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/aaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/aaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/aaa/PTCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/aaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1959	100.0	380	2	US-08-472-659-34
2	1959	100.0	380	2	US-08-474-661-34
3	1959	100.0	380	2	US-08-611-977-34
4	754.5	38.5	425	4	US-10-024-427-2
5	752	38.4	390	1	US-08-568-147B-2
6	747	38.1	390	4	US-09-266-910-3
7	745	38.0	390	4	US-09-266-910-4
8	730.5	37.3	375	1	US-08-121-714-8
9	730.5	37.3	375	1	US-08-477-108A-8
10	730.5	37.3	375	2	US-08-477-112-8
11	730.5	37.3	375	5	PCT-US93-08322-8
12	722.5	36.9	379	1	US-08-121-714-4
13	722.5	36.9	379	2	US-08-477-108A-4
14	722.5	36.9	379	2	US-08-477-112-4
15	722.5	36.9	379	5	PCT-US93-08322-4
16	715.5	36.5	391	4	US-09-123-912-110
17	715.5	36.5	391	4	US-09-643-597-110
18	715.5	36.5	391	4	US-09-480-884A-110
19	715.5	36.5	391	4	US-09-542-615A-110
20	715.5	36.5	391	4	US-09-542-615A-110
21	715	36.5	400	4	US-09-123-912-112
22	715	36.5	400	4	US-09-643-597-112
23	715	36.5	400	4	US-09-480-884A-112
24	715	36.5	400	4	US-09-542-615A-112
25	715	36.5	400	4	US-09-542-615A-112
26	673	34.4	382	1	US-07-768-286B-6
27	673	34.4	382	1	US-08-487-623B-3

28	673	34.4	382	2	US-08-997-040-3	Sequence 3, Appli
29	673	34.4	382	2	US-09-203-237-3	Sequence 3, Appli
30	671	34.3	392	1	US-07-768-286B-4	Sequence 4, Appli
31	664.5	33.9	415	1	US-07-911-531-19	Sequence 19, Appli
32	664.5	33.9	415	1	US-07-693-636A-19	Sequence 19, Appli
33	664.5	33.9	415	1	US-07-768-286B-2	Sequence 2, Appli
34	664.5	33.9	434	1	US-07-679-052A-15	Sequence 15, Appli
35	662.5	33.8	415	4	US-09-026-408-4	Sequence 2, Appli
36	662	33.8	374	1	US-08-464-148-2	Sequence 2, Appli
37	662	33.8	374	1	US-08-385-500-2	Sequence 2, Appli
38	662	33.8	374	1	US-08-846-784-2	Sequence 2, Appli
39	660.5	33.7	438	1	US-07-679-052A-17	Sequence 17, Appli
40	655	33.4	386	4	US-08-545-573A-39	Sequence 39, Appli
41	653.5	33.4	405	1	US-08-121-714-7	Sequence 7, Appli
42	653.5	33.4	405	1	US-08-477-108A-7	Sequence 7, Appli
43	653.5	33.4	405	2	US-08-477-112-7	Sequence 7, Appli
44	653.5	33.4	405	5	PCT-US93-08322-7	Sequence 7, Appli
45	652	33.3	409	4	US-09-613-303-55	Sequence 55, Appli

## ALIGNMENTS

## RESULT 1

US-08-472-659-34  
; Sequence 34, Application US/08472659  
; Patent No. 5831030  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Fuyuki  
; APPLICANT: TSUJIMOTO, No. 58310300.0  
; APPLICANT: NAKAZATO, Hiroshi  
; APPLICANT: MIURA, Kenju  
; APPLICANT: ISHIDA, No. 5831030uhiro  
; APPLICANT: KURIHARA, Tatsuya  
; APPLICANT: YAMAICHI, Kozo  
; APPLICANT: YAMAGUCHI, No. 5831030omi  
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,659  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 4-212305  
; FILING DATE: 17-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-067339  
; FILING DATE: 04-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/091,028  
; FILING DATE: 14-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McGowan, Malcolm K.  
; REGISTRATION NUMBER: 39,300  
; REFERENCE/DOCKET NUMBER: 001560-248  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 380 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-472-659-34

Query Match 100.0%; Score 1959; DB 2; Length 380;  
Best Local Similarity 100.0%; Pred. No. 1.3e-178; Indels 0; Gaps 0;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASLAANAECFNLFFEMDDNQGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60  
Db 1 MASLAANAECFNLFFEMDDNQGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60  
QY 61 NTASGYGNSNSQSGLOSKRVFSDINASHKDYDLSI VNGLFABKVGPHKDYIECAEK 120  
Db 61 NTASGYGNSNSQSGLOSKRVFSDINASHKDYDLSI VNGLFABKVGPHKDYIECAEK 120  
QY 121 LYDAKVERVDFTNHLDTRENINKWVENETHGKIKNVI GEGGISSAWVLVNAVYFKG 180  
Db 121 LYDAKVERVDFTNHLDTRENINKWVENETHGKIKNVI GEGGISSAWVLVNAVYFKG 180  
QY 181 WQSAFTKSETINCHFKSPKCGKAVAMVHQERKNLSVIEDPSMKILELRNGGINMYVL 240  
Db 181 WQSAFTKSETINCHFKSPKCGKAVAMVHQERKNLSVIEDPSMKILELRNGGINMYVL 240  
QY 241 LPENDLSEIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMKQYLRALGKDIF 300  
Db 241 LPENDLSEIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMKQYLRALGKDIF 300  
QY 301 DESKADLSGIASGGRLYISRMHKSIVIEVTEEGTEATAATGNSNIVEKQLPQSTLFRADHP 360  
Db 301 DESKADLSGIASGGRLYISRMHKSIVIEVTEEGTEATAATGNSNIVEKQLPQSTLFRADHP 360  
QY 361 FLFVIRKDDIILFSGKVSCP 380  
Db 361 FLFVIRKDDIILFSGKVSCP 380

## RESULT 2

US-08-474-661-34  
; Sequence 34, Application US/08474661  
; Patent No. 5874253  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Fuyuki  
; APPLICANT: TSURUOKA, No. 5874253uo  
; APPLICANT: NAKAZATO, Hiroshi  
; APPLICANT: MIURA, Kenju  
; APPLICANT: ISHIDA, No. 5874253uhiro  
; APPLICANT: KURIHARA, Tatsuya  
; APPLICANT: YAMAICHI, Kozo  
; APPLICANT: YAMAGUCHI, No. 5874253om.  
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: George Mason Bldg., Washington & Prince Sts.  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474,661  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/091,028

; FILING DATE: 14-JUL-1993  
; APPLICATION NUMBER: JP 4-212305  
; FILING DATE: 17-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-067339  
; FILING DATE: 04-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REA, TERESA STANKE  
; REGISTRATION NUMBER: 30,427  
; REFERENCE/DOCKET NUMBER: 001560-204  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-6620  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 380 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-474-661-34

Query Match 100.0%; Score 1959; DB 2; Length 380;  
Best Local Similarity 100.0%; Pred. No. 1.3e-178; Indels 0; Gaps 0;  
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MASLAANAECFNLFFEMDDNQGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60  
Db 1 MASLAANAECFNLFFEMDDNQGNVFFSSLSLFAALALVRLGAQDDSLSDIKLLHV 60  
QY 61 NTASGYGNSNSQSGLOSKRVFSDINASHKDYDLSI VNGLFABKVGPHKDYIECAEK 120  
Db 61 NTASGYGNSNSQSGLOSKRVFSDINASHKDYDLSI VNGLFABKVGPHKDYIECAEK 120  
QY 121 LYDAKVERVDFTNHLDTRENINKWVENETHGKIKNVI GEGGISSAWVLVNAVYFKG 180  
Db 121 LYDAKVERVDFTNHLDTRENINKWVENETHGKIKNVI GEGGISSAWVLVNAVYFKG 180  
QY 181 WQSAFTKSETINCHFKSPKCGKAVAMVHQERKNLSVIEDPSMKILELRNGGINMYVL 240  
Db 181 WQSAFTKSETINCHFKSPKCGKAVAMVHQERKNLSVIEDPSMKILELRNGGINMYVL 240  
QY 241 LPENDLSEIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMKQYLRALGKDIF 300  
Db 241 LPENDLSEIENKLTFOQLMEWTNPRMTSKYVEVFPQPKIEKNYEMKQYLRALGKDIF 300  
QY 301 DESKADLSGIASGGRLYISRMHKSIVIEVTEEGTEATAATGNSNIVEKQLPQSTLFRADHP 360  
Db 301 DESKADLSGIASGGRLYISRMHKSIVIEVTEEGTEATAATGNSNIVEKQLPQSTLFRADHP 360  
QY 361 FLFVIRKDDIILFSGKVSCP 380  
Db 361 FLFVIRKDDIILFSGKVSCP 380

## RESULT 3

US-08-611-977-34  
; Sequence 34, Application US/08611977  
; Patent No. 5972886  
; GENERAL INFORMATION:  
; APPLICANT: TSUJIMOTO, Masafumi  
; APPLICANT: IWASA, Fuyuki  
; APPLICANT: TSURUOKA, No. 5972886uo  
; APPLICANT: NAKAZATO, Hiroshi  
; APPLICANT: MIURA, Kenju  
; APPLICANT: KURIHARA, Tatsuya  
; APPLICANT: YAMAICHI, Kozo  
; APPLICANT: YAMAGUCHI, No. 5972886cmi  
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404

```

; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22133-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-611-977-34

Query Match 100.0%; Score 1959; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.3e-178;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAAANAECFNLFRMDNDQNGNVPFSSLSIPAALALVRLGAODDSLSQIDKLHV 60
DB 1 MASLAAANAECFNLFRMDNDQNGNVPFSSLSIPAALALVRLGAODDSLSQIDKLHV 60
QY 61 NTASGYGNSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYGFPHKDYIECAEK 120
DB 61 NTASGYGNSNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYGFPHKDYIECAEK 120
QY 121 LYDAKVERVDTNHLDETRNINKNWVENETHGKIKNVIGEGGSSSAVMVLVNAVYFKGK 180
DB 121 LYDAKVERVDTNHLDETRNINKNWVENETHGKIKNVIGEGGSSSAVMVLVNAVYFKGK 180
QY 181 WQSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSPMKILELRVNGGINMYVL 240
DB 181 WQSAFTKSETINCHFKSPKCGKAVAMHQRKFNLSVIEDPSPMKILELRVNGGINMYVL 240
QY 241 LPENDLSIENKLTFTQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRAALGLKIDIF 300
DB 241 LPENDLSIENKLTFTQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRAALGLKIDIF 300
QY 301 DESKADLSGIASGRRLYISRMHKSYLEVTEEGTEATAATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRRLYISRMHKSYLEVTEEGTEATAATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 4
US-10-024-427-2
; Sequence 2, Application US/10024427

; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22133-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-611-977-34

Query Match 38.5%; Score 754.5; DB 4; Length 425;
Best Local Similarity 37.4%; Pred. No. 1.4e-63;
Matches 159; Conservative 87; Mismatches 134; Indels 45; Gaps 5;

QY 1 MASLAAANAECFNLFRMDNDQNGNVPFSSLSIPAALALVRLGAODDSLSQIDKLHV 60
DB 1 MDSLVTANTKCFDLFOEIGKDDRHKNIFFSPLSLGAALGMVRLGARDSAQAQIDEVLHF 60
QY 61 NTAS-----GYGNSNSQSGLSQQLKRVF 84
DB 61 NEFSQNEKBPDPCLKSNKOKVLADSSLEGQKKTTELPQQAGSLNNSGLVSCYFGQLL 120
QY 85 SDINASHKDYDLSIVNGLFAEKYGFPHKDYIECAEKLYDAKVERVDTNHLDETRNINKN 144
DB 121 SKLDRIKTDYTLISIANELVGEQEFPCQICQYLDGVIQFYHTTIESVDFOKNPEKSRQEIF 180
QY 145 WVENETHGKIKNVIGEGGSSSAVMVLVNAVYFKGKQSAFTKSETINCHFKSPKCGKA 204
DB 181 WVECSQSGKIKELFSKDAINAETVLVNAVYFKAKWETVPDVENTVDAPFCLNANENKS 240
QY 205 VAMHQRKFNLSVIEDPSPMKILELRV-NGGINMYVLLEND-----LSIENKLTFTQ 257
DB 241 VKMTQKGLYIGFIEVRAQILEMRYTKGLSWFLPSHKNLKLGLBELEKRTIYBK 300
QY 258 LMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRAALGLKIDIFDESKADLSGIASGRRLY 317
DB 301 VWAWSSSENNSQSGLSQQLKRVFSDINASHKDYDLSIVNGLFAEKYGFPHKDYIECAEK 360
QY 318 ISRMHKSYLEVTEEGTEATAATGNSNIVEKQLPQSTLFRADHPFLFVIR-KDDIILFSG 375
DB 361 LSKTIHKTFVEVDENGTOAAATGAWSERSLSRSWVEFNANHPFLFTRHNKTQTLFYG 420
QY 376 KVSCP 380
DB 421 RVCSP 425

RESULT 5
US-08-568-147B-2
; Sequence 2, Application US/08568147B
; Patent No. 5763422
; GENERAL INFORMATION:
; APPLICANT: Suminami, Yoshinori
; APPLICANT: Kato, Hiroshi
; APPLICANT: Sekiguchi, Kiyoshi
; APPLICANT: Takeda, Katsumichi
; TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL
; TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN
; CORRESPONDENCE ADDRESS:
; ADDRESS: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York

```

```

/ COUNTRY: USA
/ ZIP: 11530
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/568,147B
/ FILING DATE:
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 800,952
/ FILING DATE: 02-DEC-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Digiglio, Frank S.
/ REGISTRATION NUMBER: 31,346
/ REFERENCE/DOCKET NUMBER: 8425
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 516-742-4343
/ TELEFAX: 516-742-4366
/ TELEX: 230 901 SANS UR
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 390 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-568-147B-2

```

```

Query Match      38.4%; Score 752; DB 1; Length 390;
Best Local Similarity 41.5%; Pred. No. 2e-63;
Matches 165; Conservative 81; Mismatches 126; Indels 26; Gaps 10;

QY 1 MASLAANAECFCNLPREDDNQNGNVFFSSLSFAALALVRLGAODDSLSQIDKLHV 60
DB 1 MNSLSEANTKFMFLFQPKRSKEN-NIFYSPISITISALGMVLLGAKNTAQIKKVLHF 59

QY 61 -----NTASGYGNSNSQSG-LQSOLKRVFSDINASHKDYDLSIUNGSLFAEKVYGFHKDY 114
DB 60 DQVTEHTTGAATYHVRSGNVHHQFKLLTEFNKSTDAVELKIANKLFGKTYLFLQY 119

QY 115 IECABKLDKAVKRVDPFTHLEDTRNINKWVENETHGKIKNVIGEGISSAVWLVA 174
DB 120 LDAIKKFTQTSVESVDFANAPESPKKINSWVSQTNKIKNLIPEGNISNTLVLVA 179

QY 175 VYFKGWOSAFKTSKTEINCHFKSPKCSKAVAMMHOERKFNLSVIEDPSMKILELRNG- 233
DB 180 IYFKQWKKKFKEDTKKEKFWPNKNTYKSIQMRQYTSFHPASLEDVQAKVLEIPYKKG 239

QY 234 GINMYVLLPE--NDLSEIENKLTQNLMEWTPNPRMTSKYVEVFFQPKIEKNYEMKOYL 291
DB 240 DUSMVLVLPNEIDGLQKLEKUTAEKLMWTSIQNNRETRVDLHLPFKVBSYDLKDTL 299

QY 292 RALGKLDIPDESADLSGIASGRLYISRMHKSXYIEVTEGTATAT-----GSNIVE 346
DB 300 RTMGWVDIFN-GDADLSGMTSGRLVSGVLHKAFAVETEGAGAAAAATAVVGSS--- 355

QY 347 KQLPOST--LFRADHPFLVIR--KDDIILFSGKVSCP 380
DB 356 ---PASTNEEHCHHPFLFFIQNKNTNSILFYGRFSSP 390

```

```

RESULT 6
US-09-266-910-3
/ Sequence 3, Application US/09266910
/ Patent No. 6344362
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: Use of a recombinant protein as receptor of a
/ TITLE OF INVENTION: hepatitis virus
/ NUMBER OF SEQUENCES: 8
/ COMPUTER READABLE FORM:

```

```

/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/266,910
/ FILING DATE:
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 390 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ US-09-266-910-3

Query Match      38.1%; Score 747; DB 4; Length 390;
Best Local Similarity 41.2%; Pred. No. 6.1e-63;
Matches 164; Conservative 81; Mismatches 127; Indels 26; Gaps 10;

QY 1 MASLAANAECFCNLPREDDNQNGNVFFSSLSFAALALVRLGAODDSLSQIDKLHV 60
DB 1 MNSLSEANTKFMFLFQPKRSKEN-NIFYSPISITISALGMVLLGAKNTAQIKKVLHF 59

QY 61 -----NTASGYGNSNSQSG-LQSOLKRVFSDINASHKDYDLSIUNGSLFAEKVYGFHKDY 114
DB 60 DQVTEHTTGAATYHVRSGNVHHQFKLLTEFNKSTDAVELKIANKLFGKTYLFLQY 119

QY 115 IECABKLDKAVKRVDPFTHLEDTRNINKWVENETHGKIKNVIGEGISSAVWLVA 174
DB 120 LDAIKKFTQTSVESVDFANAPESPKKINSWVSQTNKIKNLIPEGNISNTLVLVA 179

QY 175 VYFKGWOSAFKTSKTEINCHFKSPKCSKAVAMMHOERKFNLSVIEDPSMKILELRNG- 233
DB 180 IYFKQWKKKFKEDTKKEKFWPNKNTYKSIQMRQYTSFHPASLEDVQAKVLEIPYKKG 239

QY 234 GINMYVLLPE--NDLSEIENKLTQNLMEWTPNPRMTSKYVEVFFQPKIEKNYEMKOYL 291
DB 240 DLSMVLVLPNEIDGLQKLEKUTAEKLMWTSIQNNRETRVDLHLPFKVBSYDLKDTL 299

QY 292 RALGKLDIPDESADLSGIASGRLYISRMHKSXYIEVTEGTATAT-----GSNIVE 346
DB 300 RTMGWVDIFN-GDADLSGMTSGRLVSGVLHKAFAVETEGAGAAAAATAVVGSS--- 355

QY 347 KQLPOST--LFRADHPFLVIR--KDDIILFSGKVSCP 380
DB 356 ---PASTNEEHCHHPFLFFIQNKNTNSILFYGRFSSP 390

```

```

RESULT 7
US-09-266-910-4
/ Sequence 4, Application US/09266910
/ Patent No. 6344362
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: Use of a recombinant protein as receptor of a
/ TITLE OF INVENTION: hepatitis virus
/ NUMBER OF SEQUENCES: 8
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/266,910
/ FILING DATE:
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 390 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single

```

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
;
US-09-266-910-4

Query Match      38.0%; Score 745; DB 4; Length 390;
Best Local Similarity 41.2%; Pred. No. 9.5e-63;
Matches 164; Conservative 81; Mismatches 127; Indels 26; Gaps 10;

QY 1 MASIAAANAEEFCNLFREMDNQCNGNVPFSSLSLFAALALVRLGAQDSSLQIDKLLHV 60
Db 1 MNSLSAANTKTFDFDQOFRKSEN-NIFYSPISIIISALGMVLLGAKNTAQIKKVLHF 59

QY 61 -----NTASGYGNSNSQSG-LOSQKRVFSDINASHKDYDLSIVNGLPARKYVGHKDY 114
Db 60 DQVTEMTCKANTHIVDRSGNVHHQFOKLLTEFNKSDAYELKIANKULFGKTYLFLQY 119

QY 115 IBOAEKLYDAKVERDFTNHLNTRNINKWVENETHGKIKVIGEGISSAVMVLVNA 174
Db 120 LDAIKKFYTSVESVDFANAPESPCKINSWVESQTNKIKLIPGNGISNTIVLUNA 179

QY 175 VYKGGWQSAETKSETINCHFKSPKSGKAVAMHQBQRPNLSVIEDPSMKILELYNG- 233
Db 180 IYFKGWQEKFKSDTKEEFKFNKNTYKSIQMMRQVTSFHPASLEDVQAKVLEIDYK 239

QY 234 GINMYVLLPE--NOLSEIENKLTQNLMEWTNPRMTSKYVEVFPQPKIEKYNEMKQVL 291
Db 240 DLSMIVLLENIDGLKLEKLTAKMLNWTSLQWRETRVDLHLPRFKVEESYDLKOTL 299

QY 292 RAUGLKDIPDSKADLSGIASGRRLYISRMHKSXYIEVTEGTEATAAT-----GSNIVE 346
Db 300 RTMGWVDIFN-GDADLSGTMGRGLVSGVLHKAFFVEVTEEGARAATAVAFSS--- 355

QY 347 KQIPOST--LFRADHPPLFVIR--KDDILFSGKVCSP 380
Db 356 ---PSTNEEFHCNHPFFPIRQNTKNSILFYGRFSPP 390

RESULT 8
US-08-121-714-8
; Sequence 8, Application US/08121714
; Patent No. 5470970
;
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/121,714
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819

```

```

;
; REFERENCE/DOCKET NUMBER: 00530/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-121-714-8

Query Match      37.3%; Score 730.5; DB 1; Length 375;
Best Local Similarity 37.9%; Pred. No. 2.2e-61;
Matches 146; Conservative 85; Mismatches 135; Indels 19; Gaps 4;

QY 1 MASIAAANAEEFCNLFREMDNQCNGNVPFSSLSLFAALALVRLGAQDSSLQIDKLLHV 60
Db 1 MEQJSTANTHFADVDFRALNESDPTGNIIFISLSISSALAMIFLGTGNTAAQVSKALYF 60

QY 61 NTASGYGNSNSQSGLOSQKRVFSDINASHKDYDLSIVNGLPARKYVGHKDYIECAEK 120
Db 61 DTV-----EDHSRQSLNADINKPGAPILKLANRLTGEKTYNFLADFLASTQK 110

QY 121 LYDAKVERDFTNHLNTRNINKWVENETHGKIKVIGEGISSAVMVLVNAVYFKGK 180
Db 111 MYGAELASVDFQCAPEDARKNEINWVGQTEKIPPELLVKGVDNMTKLVNLVNAIYFGN 170

QY 181 WQSAFTKSTINCHFKSPKSGKAVAMHQBQRPNLSVIEDPSMKILELYNG-GINMYV 239
Db 171 WQEFKRETRDAPRLNKKDTKTVMQYQKCKFPYNYIEDLKCRVLELPYQKELSMII 230

QY 240 LLPEN-----DLSEIENKLTQNLMEWTNPRMTSKYVEVFPQPKIEKYNEMKQVLA 293
Db 231 LLPDIEDDESTGLEKLEKQJLTLEKREWTKPNLYLAENVVHLPFKLEESYDITSELA 290

QY 294 LGUKDIPDSKADLSGIASGRRLYISRMHKSXYIEVTEGTEATAATGSNIVEKQLPOST 353
Db 291 LGVQDLFNKRGADLSGSGARDLFVSKTIHKSFVDLNEEGTEAAAAATAGTITLAMLPEE 350

QY 354 LFRADHPPLFVIRKDDI--LFPQK 376
Db 351 NFNADHPPIFFIRNPSANILFLGR 375

RESULT 9
US-08-477-108A-8
; Sequence 8, Application US/08477108A
; Patent No. 5801001
;
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Zou, Zhigang
; APPLICANT: Anisowicz, Anthony
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,108A
; FILING DATE: June 7, 1995
; CLASSIFICATION: 536

```



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/121,714
; FILING DATE: 09/01/93
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06570/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-477-108A-8

```

```

Query Match      37.3%; Score 730.5; DB 1; Length 375;
Best Local Similarity 37.9%; Pred. No. 2.2e-61;
Matches 146; Conservative 85; Mismatches 135; Indels 19; Gaps 4;

QY 1 MASLAAANAECFNLPREMDNQNGNVFFSSLSLFAALALVRLGAGDQSLQIDKLHV 60
Db 1 MEQLSTANTHFAVDLFRALNESDPTGNIFISPLSISALAMIFLGTGRNTAAQVSKALYF 60

QY 61 NTASGYNSSNSQSGLOSKRVFSDINASHKDYDLSIVNGLFAEKVYGHKDYIECAEK 120
Db 61 DTV-----BDIHSRFQSLNADINKPGAPYILKLANRLYGEKTYNFLADFLASTOK 110

QY 121 LYDAKVERVDFTNHLDTNRNINKWVENETHGKIKVIGEGGISSAVMLVNAVYFKG 180
QY 171 WQEKWKATDAPFRINKDQTKVQMYQKKFPNYIEDLKRVLELPYQKLSMLI 230
QY 240 LIPEN-----DLSEIKLTFQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRA 293
QY 294 LGLKIDFESKADLSGIASGRLYISRMHKSIVYETEGTEATAATGNSNIVEKQLPOST 353
QY 354 LFRADHPFLFVIRKDDI--ILFSKG 376
Db 351 NFVADHPFIIFIRNPSANILFLGR 375

```

## RESULT 10

```

US-08-477-112-8
; Sequence 8, Application US/08477112
; Patent No. 5905023
; GENERAL INFORMATION:
; APPLICANT: Seger, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; City: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.

```

```

; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,112
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/121,714
; FILING DATE: 09/01/93
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06570/002003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-477-112-8

Query Match      37.3%; Score 730.5; DB 2; Length 375;
Best Local Similarity 37.9%; Pred. No. 2.2e-61;
Matches 146; Conservative 95; Mismatches 135; Indels 19; Gaps 4;

QY 1 MASLAAANAECFNLPREMDNQNGNVFFSSLSLFAALALVRLGAGDQSLQIDKLHV 60
Db 1 MEQLSTANTHFAVDLFRALNESDPTGNIFISPLSISALAMIFLGTGRNTAAQVSKALYF 60

QY 61 NTASGYNSSNSQSGLOSKRVFSDINASHKDYDLSIVNGLFAEKVYGHKDYIECAEK 120
Db 61 DTV-----BDIHSRFQSLNADINKPGAPYILKLANRLYGEKTYNFLADFLASTOK 110

QY 121 LYDAKVERVDFTNHLDTNRNINKWVENETHGKIKVIGEGGISSAVMLVNAVYFKG 180
Db 111 MYGAEIASVDFOQAPEDARKEINENWYKQTEGKIPPELLVKGVMNDMTKLVNATYFKGN 170

QY 181 WQSAFTKSETINCHFKSPKSGKAVAMMHQERKFNLSVIEDPSMKILELRYNG-GINMYV 239
Db 171 WQEKWKATDAPFRINKDQTKVQMYQKKFPNYIEDLKRVLELPYQKLSMLI 230
QY 240 LIPEN-----DLSEIKLTFQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRA 293
Db 231 LIPDDIEDESTGLEKIEKQITLEKLRWTKPENLYLAENVNHLPRFKLEESYDLTSHLAR 290

QY 294 LGLKIDFESKADLSGIASGRLYISRMHKSIVYETEGTEATAATGNSNIVEKQLPOST 353
Db 291 LGVQDLFNRGADLSGMSGARDLFVSKIIHKSFVDLNEEGTEAAAAATAGTMIAMLMPEE 350

QY 354 LFRADHPFLFVIRKDDI--ILFSKG 376
Db 351 NFVADHPFIIFIRNPSANILFLGR 375

```

## RESULT 11

```

PCT-US93-08322-8
; Sequence 8, Application PC/TUS9308322
; GENERAL INFORMATION:
; APPLICANT: Seger, Ruth

```

/ TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY

/ NUMBER OF SEQUENCES: 8

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Fish & Richardson

/ STREET: 225 Franklin Street

/ CITY: Boston

/ STATE: Massachusetts

/ COUNTRY: U.S.A.

/ ZIP: 02110-2804

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

/ COMPUTER: IBM PS/2 Model 502 or 55SX

/ OPERATING SYSTEM: MS-DOS (Version 5.0)

/ SOFTWARE: WordPerfect (Version 5.1)

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: PCT/US93/08322

/ FILING DATE:

/ CLASSIFICATION:

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: 07/938,823

/ FILING DATE: 09/01/92

/ APPLICATION NUMBER: 07/844,296

/ FILING DATE: 02/28/92

/ APPLICATION NUMBER: 07/662,216

/ FILING DATE: 02/28/91

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Fraser, Janis K.

/ REGISTRATION NUMBER: 34,819

/ REFERENCE/DOCKET NUMBER: 00530/072001

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (617) 542-5070

/ TELEFAX: (617) 542-8906

/ TELEX: 200154

/ INFORMATION FOR SEQ ID NO: 8:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 375

/ TYPE: amino acid

/ STRANDEDNESS:

/ TOPOLOGY: linear

/ PCT-US93-08322-8

Query Match

Best Local Similarity 37.3%; Score 730.5; DB 5; Length 375;

Matches 146; Conservative 85; Mismatches 135; Indels 19; Gaps 4;

QY 1 MASLAAAEFCNLEFEMDDNQGNVFPSSLSLPAALVRLGAQDLSQIDKLHV 60

Db 1 MEQLSTATHFADLFRALNESDPTGNIFSPISLSALAMIFLGTGNTAAQVSKALYF 60

QY 61 NTASGYGNSNSQSGLSQKRVFSDINASHKDYDLSIVNGLFAKVVGFHKDYIECAEK 120

Db 61 DTV-----EDHSRFQSLNADINKGASVILKANRLYGEKTYNPLADFLASTQK 110

QY 121 LYDAKVEVDPTNHLIEDTRNINKVNEETHGKIKNVIGEGGSISSAWVLNVAVYFKG 180

Db 111 MYGAELASVDFQAPEDARKEINWVKQTEGKIPELLVKGWDMWTKLVNLYAFKGN 170

QY 181 WSAFTKSETINCHFKSPCKGKAVAMWQERFENLSVIEDPSMKILERYNG-GINMYV 239

Db 171 WEKFMKATRDAPFLNKKOTKTVRMVYQKXFPYNYIEDLKRCVLELPYQGEELSMII 230

QY 240 LLPEN-----DLSEIENKLTQNLMEWTNPRMTSKVVEFFPQFKIEKNYEMQYLRA 293

Db 231 LLPDDIEDSTGLKIEEQLTLEKLEHWTPENLYIAEWNHLFRFKLESVDLTSILAR 290

QY 294 LGLKIDFESKADLSAGGLYISRMWHSKYVIEVTEGTEATATGNSNIVEKOLPOST 353

Db 291 LGVQDLFRNGKADLSGARDLFVSKIIHKSVDLNEEGTEGTAATACTIMLANLPEE 350

QY 354 LFRADHPFLVIRKDDI--ILFSQK 376

Db 351 NFNADHPFIFFIRHNPANSNIILFGR 375

RESULT 12

US-08-121-714-4

/ Sequence 4, Application US/08121714

/ Patent No. 5470970

/ GENERAL INFORMATION:

/ APPLICANT: Sager, Ruth

/ TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH

/ TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY

/ NUMBER OF SEQUENCES: 8

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Fish & Richardson

/ STREET: 225 Franklin Street

/ CITY: Boston

/ STATE: Massachusetts

/ COUNTRY: U.S.A.

/ ZIP: 02110-2804

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

/ COMPUTER: IBM PS/2 Model 502 or 55SX

/ OPERATING SYSTEM: MS-DOS (Version 5.0)

/ SOFTWARE: WordPerfect (Version 5.1)

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/121,714

/ FILING DATE:

/ CLASSIFICATION: 536

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: 07/938,823

/ FILING DATE: 09/01/92

/ APPLICATION NUMBER: 07/844,296

/ FILING DATE: 02/28/92

/ APPLICATION NUMBER: 07/662,216

/ FILING DATE: 02/28/91

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Fraser, Janis K.

/ REGISTRATION NUMBER: 34,819

/ REFERENCE/DOCKET NUMBER: 00530/072001

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (617) 542-5070

/ TELEFAX: (617) 542-8906

/ TELEX: 200154

/ INFORMATION FOR SEQ ID NO: 4:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 379

/ TYPE: amino acid

/ STRANDEDNESS:

/ TOPOLOGY: linear

/ US-08-121-714-4

Query Match 36.9%; Score 722.5; DB 1; Length 379;

Best Local Similarity 38.6%; Pred. No. 1.3e-60;

Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

QY 1 MASLAAAEFCNLEFEMDDNQGNVFPSSLSLPAALVRLGAQDLSQIDKLHV 60

Db 1 MEQLSANTRALDFALESENPNAGNIFSPISLSAMAVFLGTGNTAAQLSTHFF 60

QY 61 NTASGYGNSNSQSGLSQKRVFSDINASHKDYDLSIVNGLFAKVVGFHKDYIECAEK 120

Db 61 NTV-----EVSRSFQSLNADINKGASVILKANRLYGEKTYNPLFELVSTQK 110

QY 121 LYDAKVEVDPTNHLIEDTRNINKVNEETHGKIKNVIGEGGSISSAWVLNVAVYFKG 180

Db 111 TYGADLASVDFOASDARKEINWVKQTEGKIPELLASGWDNMTKLVNLYAFKGN 170

QY 181 WSAFTKSETINCHFKSPCKGKAVAMWQERFENLSVIEDPSMKILERYNG-GINMYV 239

Db 171 WDKFMKATRDAPFLNKKOTKTVRMVYQKXFPYNYIEDLKRCVLELPYQGEELSMVI 230

QY 240 LLPEN-----DLSEIENKLTQNLMEWTNPRMTSKVVEFFPQFKIEKNYEMQYLRA 293

Db 231 LLPDDIEDSTGLKIEEQLTLEKLEHWTPENLYIAEWNHLFRFKLESVDLTSILAR 290

QY 294 LGLKDFDESKADLSGASGRILYSRMWHSYIEVTEGTEATAATGSSNIVEKQLPOST 353  
 Db 291 LGVQDLFNSKADLSGSGARDIFISKIVHKSFEVNEEGTEAAAAAGIATGFCMLPPE 350  
 QY 354 LFRADHPFLFVIRKDD--IILFSGKVSCP 380  
 Db 351 NPTADHPFLFIRNSSGSIILFLGRFSSP 379

## RESULT 13

US-08-477-108A-4  
 ; Sequence 4, Application US/08477108A  
 ; Patent No. 5801001  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sager, Ruth  
 ; APPLICANT: Zou, Zhigang  
 ; APPLICANT: Anisowicz, Anthony  
 ; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH  
 ; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTER: IBM PS/2 Model 502 or 55SX  
 OPERATING SYSTEM: MS-DOS (Version 5.0)  
 SOFTWARE: WordPerfect (Version 5.1)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/477,108A  
 FILING DATE: June 7, 1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/121,714  
 FILING DATE: 09/01/93  
 APPLICATION NUMBER: 07/938,823  
 FILING DATE: 09/01/92  
 APPLICATION NUMBER: 07/844,296  
 FILING DATE: 02/28/92  
 APPLICATION NUMBER: 07/662,216  
 FILING DATE: 02/28/91  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fraser, Janis K.  
 REGISTRATION NUMBER: 34,819  
 REFERENCE/DOCKET NUMBER: 06570/002002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-5070  
 TELEFAX: (617) 542-8906  
 TELEX: 200154  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 379  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear

US-08-477-108A-4

Query Match 36.9%; Score 722.5; DB 1; Length 379;  
 Best Local Similarity 38.6%; Pred. No. 1.3e-60;  
 Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

QY 1 MASLAANAECFNLFFEMDQNGNVPFSSLSLFAALVRLGAGDDSLSDKLLHV 60  
 Db 1 MEQLSANTRALDLFLALSNPAGNIFISFSSISAMAVFLGTRGNTAQLSKTHF 60  
 QY 61 NTASGYGNSNSQSGLSQKRVFSDINASHKDYDLSVNGLFPAKVGTHGYIECEK 120  
 Db 61 NTV-----BEVHSRFQSNADINKRGASYILKANRLYGEKTYNLPFLVSTQK 110

QY 121 LYDAKVERVDFTNHLERTRNINKWENETHGKIKNIVGEGISSSAWVLVNAVIFKQK 180  
 Db 111 TYGADLASVDYQHASDEARKTNQWVKGQTEGKIPELLASGMDVDMTKLVNAVIFKGN 170  
 QY 181 WQSAFTKSTETINCHFKSPKSGKAVAMMHQBRKXENLSVIEDSPSKILRLRYNG-GINMYV 239  
 Db 171 WKQKFKKEATNAPFLNKDKRKYVMYQKKKPAFYGYIEDLKCRVLELPYQGEELSMVI 230  
 QY 240 LLPEN-----DLSEIENKLTQNLMEWTNPRMTSKYVEVFPQFKIEKNYMKQYLR 293  
 Db 231 LLPDDIEDBSTGLKKIEEQTLLEKHEWTKPENLDPIEVNLSLPRFKLEBSYTLNSDLAR 290  
 QY 294 LGLKDFDESKADLSGASGRILYSRMWHSYIEVTEGTEATAATGSSNIVEKQLPOST 353  
 Db 291 LGVQDLFNSKADLSGSGARDIFISKIVHKSFEVNEEGTEAAAAAGIATGFCMLPPE 350  
 QY 354 LFRADHPFLFVIRKDD--IILFSGKVSCP 380  
 Db 351 NPTADHPFLFIRNSSGSIILFLGRFSSP 379

## RESULT 14

US-08-477-112-4  
 ; Sequence 4, Application US/08477112  
 ; Patent No. 5905023  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sager, Ruth  
 ; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH  
 ; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTER: IBM PS/2 Model 502 or 55SX  
 OPERATING SYSTEM: MS-DOS (Version 5.0)  
 SOFTWARE: WordPerfect (Version 5.1)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/477,112  
 FILING DATE:  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/121,714  
 FILING DATE: 09/01/93  
 APPLICATION NUMBER: 07/938,823  
 FILING DATE: 09/01/92  
 APPLICATION NUMBER: 07/844,296  
 FILING DATE: 02/28/92  
 APPLICATION NUMBER: 07/662,216  
 FILING DATE: 02/28/91  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fraser, Janis K.  
 REGISTRATION NUMBER: 34,819  
 REFERENCE/DOCKET NUMBER: 06570/002003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-5070  
 TELEFAX: (617) 542-8906  
 TELEX: 200154  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 379  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear

US-08-477-112-4

Query Match 36.9%; Score 722.5; DB 2; Length 379;  
 Best Local Similarity 38.6%; Pred. No. 1.3e-60;

Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

```

QY 1 MASLAANAACFCNLFREMDNQGNGVFFSSLSLFAALALVRLGAQDDSLQIDKILHV 60
Db 1 MEQSSANTFAIDFLALSENPNAGNIFISPFSSSAWAVFLGTRGNTAAQLSKTHF 60
QY 61 NTASGYGNSNSQSLQKRVFSDINASHKDYDLISVINGLFAEKVYGFHDKYTECAEK 120
Db 61 NTV-----EEVHSRFQSLNADINKGASYILKLANLYGKTYNLFPEFLVSTQK 110
QY 121 LYDAKVERVDFTNHLSDTRRNINKNWENETHGKINNVIGEGGSSSAVMVLVNAVYFKGK 180
Db 111 TYGADLASVDFOHASEDARKTINQWKGTEGKIPELLASGVDNMTKILVNAVYFKGK 170
QY 181 WQSAFTKSETINCHPKSCGKAVAMHOKERKNLSVIEDPSMKILELYNG-GINMYV 239
Db 171 WKDFMKEATTNAPFLNKKDKRKTVKMYCKKFAYGIEDLRCRVLELPYQGBELSMVI 230
QY 240 LLPEN-----DLSEIENKLTFOQLMEWTNPRMTSKYVEVFPFQPKIEKMYEMQVIRA 293
Db 231 LLPDDIDESGTGLKIEEQLTLEKLEHWTKPENLDFIEVNVSLPRFKLEESTYTLNSDLAR 290
QY 294 LGLKDIPEKADSLASGGRLYISMMHKSIVTEBTEGTEATAATGNSNIVERKQLPOST 353
Db 291 LGVQDLFNSGADLSGMSGARDIFISKIVHKSFEVNEEGTEAAATAGIATFCMLMPEE 350
QY 354 LFRADHPFLFVIRKDD--IILFSGKYSCP 380
Db 351 NFTAADHPFLFFIRHNSGSGILFLGRFSSP 379

```

## RESULT 15

```

PCT-US93-08322-4
; Sequence 4, Application PC/TUS9308322
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08322
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/26/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379
; TYPE: amino acid

```

```

; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-08322-4

```

```

Query Match 36.9%; Score 722.5; DB 5; Length 379;
Best Local Similarity 38.6%; Pred. No. 1.3e-60;
Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

```

```

QY 1 MASLAANAACFCNLFREMDNQGNGVFFSSLSLFAALALVRLGAQDDSLQIDKILHV 60
Db 1 MEQSSANTFAIDFLALSENPNAGNIFISPFSSSAWAVFLGTRGNTAAQLSKTHF 60
QY 61 NTASGYGNSNSQSLQKRVFSDINASHKDYDLISVINGLFAEKVYGFHDKYTECAEK 120
Db 61 NTV-----EEVHSRFQSLNADINKGASYILKLANLYGKTYNLFPEFLVSTQK 110
QY 121 LYDAKVERVDFTNHLSDTRRNINKNWENETHGKINNVIGEGGSSSAVMVLVNAVYFKGK 180
Db 111 TYGADLASVDFOHASEDARKTINQWKGTEGKIPELLASGVDNMTKILVNAVYFKGK 170
QY 181 WQSAFTKSETINCHPKSCGKAVAMHOKERKNLSVIEDPSMKILELYNG-GINMYV 239
Db 171 WKDFMKEATTNAPFLNKKDKRKTVKMYCKKFAYGIEDLRCRVLELPYQGBELSMVI 230
QY 240 LLPEN-----DLSEIENKLTFOQLMEWTNPRMTSKYVEVFPFQPKIEKMYEMQVIRA 293
Db 231 LLPDDIDESGTGLKIEEQLTLEKLEHWTKPENLDFIEVNVSLPRFKLEESTYTLNSDLAR 290
QY 294 LGLKDIPEKADSLASGGRLYISMMHKSIVTEBTEGTEATAATGNSNIVERKQLPOST 353
Db 291 LGVQDLFNSGADLSGMSGARDIFISKIVHKSFEVNEEGTEAAATAGIATFCMLMPEE 350
QY 354 LFRADHPFLFVIRKDD--IILFSGKYSCP 380
Db 351 NFTAADHPFLFFIRHNSGSGILFLGRFSSP 379

```

```

Search completed: December 12, 2003, 16:20:06
Job time : 23 secs

```



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-140-719-34

```

```

Query Match 100.0%; Score 1959; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 8e-172;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MASLAANAFCNLFREDDNQNGNVFSSLSLFAALALVRLGAQDDSLQIDKLLHV 60
DB 1 MASLAANAFCNLFREDDNQNGNVFSSLSLFAALALVRLGAQDDSLQIDKLLHV 60
QY 61 NTASGYGNSNSOSGLQSLKRVFSDINASHKDYDLSTVNGLPAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSOSGLQSLKRVFSDINASHKDYDLSTVNGLPAEKVYGFHKDYIECAEK 120
QY 121 LYDAKVRVDFTHLEDDTRNINKNWENETHGKIKNVIGEGGISSAVMVLNAVYFKG 180
DB 121 LYDAKVRVDFTHLEDDTRNINKNWENETHGKIKNVIGEGGISSAVMVLNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSKAVAMMHOERKPNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSKAVAMMHOERKPNLSVIEDPSMKILELYNGGINMYVL 240
QY 241 LPENDLSEIENKLTIFQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRALGLKDF 300
DB 241 LPENDLSEIENKLTIFQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRALGLKDF 300
QY 301 DESKADLSGIASGRRLYISRMWHSYIEVTEEGTATAATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRRLYISRMWHSYIEVTEEGTATAATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

```

## RESULT 2

US-10-091-442-34

Sequence 34, Application US/10091442

Publication No. US20020164711a1

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IWASA, Fuyuki

TSURUOKA, No. US20020164711a1

NAKAZATO, Hiroshi

MIURA, Kenju

ISHIDA, No. US20020164711a1

KURIHARA, Tatsuya

YAMAUCHI, Kozo

YAMAGUCHI, No. US20020164711a1

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker &amp; Mathis

STREET: P.O. Box 1404

```

; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/091,442
; FILING DATE: 07-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

```

```

Query Match 100.0%; Score 1959; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 8e-172;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAFCNLFREDDNQNGNVFSSLSLFAALALVRLGAQDDSLQIDKLLHV 60
DB 1 MASLAANAFCNLFREDDNQNGNVFSSLSLFAALALVRLGAQDDSLQIDKLLHV 60
QY 61 NTASGYGNSNSOSGLQSLKRVFSDINASHKDYDLSTVNGLPAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSOSGLQSLKRVFSDINASHKDYDLSTVNGLPAEKVYGFHKDYIECAEK 120
QY 121 LYDAKVRVDFTHLEDDTRNINKNWENETHGKIKNVIGEGGISSAVMVLNAVYFKG 180
DB 121 LYDAKVRVDFTHLEDDTRNINKNWENETHGKIKNVIGEGGISSAVMVLNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSKAVAMMHOERKPNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSKAVAMMHOERKPNLSVIEDPSMKILELYNGGINMYVL 240
QY 241 LPENDLSEIENKLTIFQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRALGLKDF 300
DB 241 LPENDLSEIENKLTIFQNLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRALGLKDF 300
QY 301 DESKADLSGIASGRRLYISRMWHSYIEVTEEGTATAATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRRLYISRMWHSYIEVTEEGTATAATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

```



Query Match 41.8%; Score 819; DB 15; Length 392;  
Best Local Similarity 43.1%; Pred. No. 7.1e-67;  
Matches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

Qy 1 KASIAAANAFECNLFREMDNQGNGVFFSSLSLFAALALVFLGAQDDLSQIDKLLHV 60  
Db 1 MGSLSANVEFCLDVFKELSNNGNIGNIFSSLSLLVALSNVLLGARGETAQLEKVLHF 60

Qy 61 -----NTASGYNSSN--SQSG--LQSOLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHD 113  
Db 61 SHTVDSLKPGFKDSCQAGRHSFGVBFQINQPSNCTLSIANRLYGTKTMAHQ 120

Qy 114 YIECAEKLDAKVERVDTNHELETRNKNVNEHKGKIXVIGEGISSAVMYLVN 173  
Db 121 YLSCEKWOQARLQTVDFEQSTETRMNINAVENKINGVANLFGKSTIDPSSVMVLN 180

Qy 174 AVYFKGWSQATKSETINCHFKSPKCSGKAVAMHQRKPNLSVIEDPSMKILELY-N 232  
Db 181 IYIFKGRQNKFTVETKSPFQISGKNVTVMYQIGTFKLAFAVKEPQVLELYN 240

Qy 233 GGINMYVLLPEN--DLSEIENKLTQNLMEWTNPRMTSKYVEVFPFQFKIEKNYMKQV 290  
Db 241 NKLSMILLVPGIANLQKQLNSGTTFHEWTSNNMWEREVEVHLPRFKLEIKYELN 300

Qy 291 LRALGLKDIPDSKADLSGASGRGLYISRMHKSXYIEVTEGTEATAATGSNIVEKQLP 350  
Db 301 LKPLGVTLDFNQKADLSGMSPTKGLYLKAIHKSYLDVSEEGTEAATAATGDSIAVKS 360

Qy 351 QSTLFRADHPFLFVIR--KDDIILFSGKVSCLP 380  
Db 361 MEAQFKANHPFLFIRHTHTNTLFCGKLASP 392

## RESULT 6

US-09-895-814-946  
; Sequence 946, Application US/09895814  
; Publication No. US20020193296A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C26  
; CURRENT APPLICATION NUMBER: US/09/895,814  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 990  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 946  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-895-814-946

Query Match 40.7%; Score 796.5; DB 10; Length 379;  
Best Local Similarity 42.6%; Pred. No. 7.9e-65;  
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

Qy 1 KASIAAANAFECNLFREMDNQGNGVFFSSLSLFAALALVFLGAQDDLSQIDKLLHV 60  
Db 1 MGSLSANVEFCLDVFKELSNNGNIGNIFSSLSLLVALSNVLLGARGETAQLEKVLHF 60

Qy 61 -----NTASGYNSSN--SQSG--LQSOLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHD 113  
Db 61 SHTVDSLKPGFKDSCQAGRHSFGVBFQINQPSNCTLSIANRLYGTKTMAHQ 120

Qy 114 YIECAEKLDAKVERVDTNHELETRNKNVNEHKGKIXVIGEGISSAVMYLVN 173  
Db 121 YLSCEKWOQARLQTVDFEQSTETRMNINAVENKINGVANLFGKSTIDPSSVMVLN 180

Qy 174 AVYFKGWSQATKSETINCHFKSPKCSGK--AVAMHQRKPNLSVIEDPSMKILELY 231  
Db 181 IYIFKGRQNKFTVETKSPFQISGKNVTVMYQIGTFKLAFAVKEPQVLELY 225

Qy 232 -NGGINMYVLLPEN--DLSEIENKLTQNLMEWTNPRMTSKYVEVFPFQFKIEKNYEMK 288  
Db 226 YNNKLSMILLVPGIANLQKQLNSGTTFHEWTSNNMWEREVEVHLPRFKLEIKYELN 285

Qy 289 QYLRALGLKDIPDSKADLSGASGRGLYISRMHKSXYIEVTEGTEATAATGSNIVEKQ 348  
Db 286 SLKPLGVTLDFNQKADLSGMSPTKGLYLKAIHKSYLDVSEEGTEAATAATGDSIAVKS 345

Qy 349 LPQSTLFRADHPFLFVIR--KDDIILFSGKVSCLP 380  
Db 346 LPMRAQFKANHPFLFIRHTHTNTLFCGKLASP 379

## RESULT 7

US-10-144-678A-946  
; Sequence 946, Application US/10144678A  
; Publication No. US20030157089A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Deng, Ta  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C28  
; CURRENT APPLICATION NUMBER: US/10/144,678A  
; CURRENT FILING DATE: 2002-08-12  
; NUMBER OF SEQ ID NOS: 1033  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 946  
; LENGTH: 379  
; TYPE: PRT



```

; ORGANISM: Homo sapiens
US-10-144-678A-946

Query Match      40.7%; Score 796.5; DB 12; Length 379;
Best Local Similarity 42.6%; Pred. No. 7.9e-65;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 1 MASLAANAFCNLFREDDNQGNGNVFSSLSFAALALVRLGAQDLSQIDKLHV 60
Db 1 MGSUSTANVFCLDVFKELNSNIGDNIFFSSLSLYALSVMVLGARGTAEQLEKVLHF 60
QY 61 -----NTASGVGNSSN-SQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAKVVGPXKD 113
Db 61 SHTVDSLKPGFKDSPKCSQAGSIHSEGFVFSQINQPSNCTLSIANRLYGTWAFHQ 120
QY 114 YIECAEKLYDAKVERVDFTNHLETRNINKWVENETGKIKNVIGEGGSISSAVMLVN 173
Db 121 YLSCSEKWOARLQTVDFEQSTEETRMINAWENKNGKVANLFGKSTIDPSSVMVLN 180
QY 174 AVYFKGQWQSAFTKSTINCHFKSPKCSGK--AVAMHQRKFNLSVIEDPSMKILELY 231
Db 181 TIYFKGQRQNF-----QGNVTVMYQIGTFKLAFFVKPQMVLELPY 225
QY 232 -NGGINMYVLLPEN--DLSEIKNLITFQNLMEWTNPRMTSKYVEVFPFQPKIEKNYEMK 288
Db 226 VNNKLSMILLPVGIANLQIEKQNSGTHEWTSNNMEREVEVHLPRFKLEIKYELN 285
QY 289 QYLRALGLKDIFDESADLSGASGRLYISRMMHKSIVTEEGTEATAATGSNIVEKQ 348
Db 286 SLLKPLGVTDLFNQVKADLSGMSPTKGLYLSKAIHKSILYDVSSEGTAAATGDSIAVKS 345
QY 349 LPOSTLFRADHPFLVIR--KDDILFSGKVSCP 380
Db 346 LPMRAQFKANHPFLFIRHTHTNLTILFCGLASP 379

RESULT 8
US-10-294-025-946
; Sequence 946, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294.025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-946

Query Match      40.7%; Score 796.5; DB 12; Length 379;
Best Local Similarity 42.6%; Pred. No. 7.9e-65;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 1 MASLAANAFCNLFREDDNQGNGNVFSSLSFAALALVRLGAQDLSQIDKLHV 60
Db 1 MGSUSTANVFCLDVFKELNSNIGDNIFFSSLSLYALSVMVLGARGTAEQLEKVLHF 60
QY 61 -----NTASGVGNSSN-SQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAKVVGPXKD 113
Db 61 SHTVDSLKPGFKDSPKCSQAGSIHSEGFVFSQINQPSNCTLSIANRLYGTWAFHQ 120
QY 114 YIECAEKLYDAKVERVDFTNHLETRNINKWVENETGKIKNVIGEGGSISSAVMLVN 173
Db 121 YLSCSEKWOARLQTVDFEQSTEETRMINAWENKNGKVANLFGKSTIDPSSVMVLN 180
QY 174 AVYFKGQWQSAFTKSTINCHFKSPKCSGK--AVAMHQRKFNLSVIEDPSMKILELY 231
Db 181 TIYFKGQRQNF-----QGNVTVMYQIGTFKLAFFVKPQMVLELPY 225
QY 232 -NGGINMYVLLPEN--DLSEIKNLITFQNLMEWTNPRMTSKYVEVFPFQPKIEKNYEMK 288
Db 226 VNNKLSMILLPVGIANLQIEKQNSGTHEWTSNNMEREVEVHLPRFKLEIKYELN 285
QY 289 QYLRALGLKDIFDESADLSGASGRLYISRMMHKSIVTEEGTEATAATGSNIVEKQ 348
Db 286 SLLKPLGVTDLFNQVKADLSGMSPTKGLYLSKAIHKSILYDVSSEGTAAATGDSIAVKS 345
QY 349 LPOSTLFRADHPFLVIR--KDDILFSGKVSCP 380
Db 346 LPMRAQFKANHPFLFIRHTHTNLTILFCGLASP 379

RESULT 9
US-10-012-896-946
; Sequence 946, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yudi
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Keppler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Neagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-946

Query Match      40.7%; Score 796.5; DB 14; Length 379;
Best Local Similarity 42.6%; Pred. No. 7.9e-65;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 1 MASLAANAFCNLFREDDNQGNGNVFSSLSFAALALVRLGAQDLSQIDKLHV 60
Db 1 MGSUSTANVFCLDVFKELNSNIGDNIFFSSLSLYALSVMVLGARGTAEQLEKVLHF 60
QY 61 -----NTASGVGNSSN-SQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAKVVGPXKD 113
Db 61 SHTVDSLKPGFKDSPKCSQAGSIHSEGFVFSQINQPSNCTLSIANRLYGTWAFHQ 120
QY 114 YIECAEKLYDAKVERVDFTNHLETRNINKWVENETGKIKNVIGEGGSISSAVMLVN 173
Db 121 YLSCSEKWOARLQTVDFEQSTEETRMINAWENKNGKVANLFGKSTIDPSSVMVLN 180
QY 174 AVYFKGQWQSAFTKSTINCHFKSPKCSGK--AVAMHQRKFNLSVIEDPSMKILELY 231
Db 181 TIYFKGQRQNF-----QGNVTVMYQIGTFKLAFFVKPQMVLELPY 225
QY 232 -NGGINMYVLLPEN--DLSEIKNLITFQNLMEWTNPRMTSKYVEVFPFQPKIEKNYEMK 288
Db 226 VNNKLSMILLPVGIANLQIEKQNSGTHEWTSNNMEREVEVHLPRFKLEIKYELN 285
QY 289 QYLRALGLKDIFDESADLSGASGRLYISRMMHKSIVTEEGTEATAATGSNIVEKQ 348
Db 286 SLLKPLGVTDLFNQVKADLSGMSPTKGLYLSKAIHKSILYDVSSEGTAAATGDSIAVKS 345
QY 349 LPOSTLFRADHPFLVIR--KDDILFSGKVSCP 380
Db 346 LPMRAQFKANHPFLFIRHTHTNLTILFCGLASP 379
```

Db 121 YLSCSEKWKYQARLOTVDPEQSTBETRKMINAWENKTNCKVANLFCKSTIDPSSVMVLVN 180  
Qy 174 AVYEKWKQWQSAFTKSEITINCHFKSPKCSK--AVAMMEQERKFNLSVIEDPSMKILELRY 231  
Db 181 TIIFPKQQRNK?-----QCKVTVEMMYQIGFKLAFVKEPQMOVLELRY 225  
Qy 232 -NGGINMYVLLPEN--DLSEIENKLTFFQNAEWTPNPRMTSKYVFFPQFKIEKNYEMK 288  
Db 256 VNKLSMIILLPVGINLQLEKQLNSGTFHEWTSNNMEREVEVHLPRFKLEIKYELN 285  
Qy 289 QYLRALGLKDIIDFSKADLSGIASGGRLYISNMHKSYLEVTEEGTEATATGNSVNEKQ 348  
Db 286 SLIKPLGVTDLFPNQWADLSGMPFKGLYLKAIHKSYLEVDSSEGEAAATGSAVKS 345  
Qy 349 LPQSTLFRADHPFLFVIR--KDDIILFSGKVSCLP 380  
Db 346 LPKRAQKANHFFLEFIRHETWILLFCGLKASLP 379

RESULT 10  
US-09-823-187-2  
; Sequence 2, Application US/09823187  
; Publication No. US20030096952A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgesse, Catherine  
; APPLICANT: Gusev, Vladimir Y  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Padigaru, Muralidhar  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Spyttek, Kimberly  
; APPLICANT: Taupier, Raymond J  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15968-745  
; CURRENT APPLICATION NUMBER: US/09/823,187  
; CURRENT FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/193,339  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: 60/193,205  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: 60/195,343  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: 60/195,088  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: 60/195,005  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: 60/195,792  
; PRIOR FILING DATE: 2000-04-10  
; PRIOR APPLICATION NUMBER: 60/196,556  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: 60/197,081  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: 60/197,525  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/197,087  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 103  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 395  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-823-187-2

Query Match 39.4%; Score 771.5; DB 11; Length 395;  
Best Local Similarity 40.3%; Pred. No. 1.7e-62;  
Matches 159; Conservative 89; Mismatches 132; Indels 15; Gaps 5;  
Qy 1 MASLAAANAEFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLQIDKLHV 60  
Db 1 MDSLVTANTKFCDFLQEGIKGDRHKNIFFSPLSLSAALGMVRLGARSASAHQIDEVLHF 60

Qy 61 NTASGY-----GNSSNSQSGLQS-OLKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKQY 114  
Db 61 NEFQNESKEPAGSLNNESGLVSCYFQLLSKLDRIKTDYTLISIANRLYGEFFPICOEY 120  
Qy 115 IECSEKLYDAKVERVDYFNHLETRTNINKVVENETHGKIKNVIHGEGISSATWVLVNA 174  
Db 121 LDGVIQFYHTTIESYDFQKPEKSQRIENFWECOSQKIKDLFSKDAINAEIVLVLVNA 180  
Qy 175 VYFKGKQWQSAFTKSTINCHFKSPKCSKAVAMMHQERKFNLSVIDPSMKILELRY-NG 233  
Db 181 VYFAKWEIYFHEHTVDAFPCLNQENKSVQMTQKGLYRIGFIEVKAQILEMRYTKG 240  
Qy 234 GINMYVLLPEN-----LSIENKLTFFQNAEWTPNPRMTSKYVFFPQFKIEKNYEM 287  
Db 241 KLSMFVLLPSHSDNLKGLIELEKRTIYKMWANSSSENSESVVLSFPFRFTLESDYL 300  
Qy 288 QYLRALGLKDIIDFSKADLSGIASGGRLYISNMHKSYLEVTEEGTEATATGNSVNEK 347  
Db 301 NSIILQDMGITDIFDETRADLTGISPSNLYLSKIHKITFEVDENGTOAAATGAVVSR 360  
Qy 348 QLPQSTLFRADHPFLFVIR--KDDIILFSGKVSCLP 380  
Db 361 SLRSWVEFNANHPFLFIRHKNKTQTLFYGRVCSLP 395

RESULT 11  
US-10-094-944-2  
; Sequence 2, Application US/10094944  
; Publication No. US20030160275A1  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Clarke, Howard RG  
; APPLICANT: DuBoise, Robert F  
; APPLICANT: Wiley, Steven R  
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES  
; FILE REFERENCE: 3223-A  
; CURRENT APPLICATION NUMBER: US/10/094,944  
; CURRENT FILING DATE: 2002-03-08  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 425  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-944-2

Query Match 38.5%; Score 754.5; DB 12; Length 425;  
Best Local Similarity 37.4%; Pred. No. 6.9e-61;  
Matches 159; Conservative 87; Mismatches 134; Indels 45; Gaps 5;  
Qy 1 MASLAAANAEFCNLFREDDNQNGNVFFSSLSLFAALALVRLGAQDDSLQIDKLHV 60  
Db 1 MDSLVTANTKFCDFLQEGIKGDRHKNIFFSPLSLSAALGMVRLGARSASAHQIDEVLHF 60  
Qy 61 NTAS-----GFGSSNSQSGLQS-OLKRVF 84  
Db 61 NEFQNESKEPAGSLNNESGLVSCYFQLLSKLDRIKTDYTLISIANRLYGEFFPICOEY 120  
Qy 85 SDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEKLKYDAKVERVDYFNHLETRTNINK 144  
Db 121 SKLDRIKTDYTLISIANRLYGEFFPICOEYVLDGVTFYHTIESVDFQKPKSQRIENFW 180  
Qy 145 WYENETHGKIKNVIHGEGISSAYVNLVNAVYFKGQWQSAFTKSTINCHFKSPKCSKA 204  
Db 181 WYECOSQKIKELFSKDAINAEIVLVLVNAVYFKAKWTYFDHENTVDAPFCINANENKS 240  
Qy 205 VAMMEQERKFNLSVIEDPSMKILELRY-NGGINMYVLLPEN-----LSEIENKLTFFQ 257  
Db 241 VQMTQKGLYRIGFIEVKAQILEKRYTKGLSMFVLLPSHSDNLKGLIELEKRTIYK 300  
Qy 258 LMEWTPNPRMTSKYVFFPQFKIEKNYEMKQYLRALGLKDIIDFSKADLSGIASGGRLY 317

```
Db 301 MWANSSNNSESUVLSPFRFTLEDSDYDLSNLQDMGITDIFDETRADLTGISSPNLY 360
QY 318 ISRMHKSXYIEVTEGTEATAATGNSIVEXQLFQSTLFRADHPFLFVIR--KDDIILFSG 375
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 361 LSKIIHKTFFVEDNGTQAATGAUVSESSLSWSVEFNAHNPFLPIRHNKTQTILFYG 420
QY 376 KYSCP 380
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 421 RVCSP 425

RESULT 12
US-09-823-187-29
; Sequence 29, Application US/09823:187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-29

Query Match 38.4%; Score 752; DB 11; Length 390;
Best Local Similarity 41.5%; Pred. No. 1e-60;
Matches 165; Conservative 81; Mismatches 126; Indels 26; Gaps 10;
QY 1 MASLAANAECFNLPREMDNQGNGNVFFSSLSLFAALALVRLGAQDLSLQIDKLHV 60
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MNSLSEANTKFMFDLQQQRKSKEN-NIFYSPISITSALGWLGLGAKDNTAQIKKVLHF 59
QY 61 -----NTASGVGNSSNSQS-LQSLKXRVFSDINASHKDYDLSTVNCGLFAKXVYGFHKDY 114
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 60 DQVTEITGKAANYHVRDGNVHHQFQKILTEINFKSTDAYELKIANKLFGEKTYDLEQY 119
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 115 IECAKLYDAKVERVDFTNHLEDTRRNINKWENETHGKIKNVIGEGGTSAAVAVLVNA 174
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 120 LQAIKKFYQSVSDVFANAPESRKKINSWVESQTNKEIKNLIIPSGNIGSTILVLNA 179
```

```
QY 175 VVFGKQWSAFKSETINCHFKSPKCGKAVAMQHQRKFNLSVIEDPSMKILEFYNG- 233
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 180 IYFKQWEKKFNKEDIKEEKFPWPKNTVYKIQMRQYTSFHFASLEDVQAKVLEIPYKKG 239
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 234 GINMYVLLPE--NDLSIEENKLTQNLMEWTNPRMTSKYIEVFFPOPKIEKNYKQYL 291
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 240 DLSMTVLLPNEIDGLQKLEKLTAEKLMWTSIQNNVRETRVDLHLPKRVESYDKDTL 299
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 292 RALGLKQIFDSKADLSGIASGRLYISRMWHSXYIEVTEGTEATAAT-----GSNIVE 346
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 300 RTWGVVDIFN-GDADLSGMTGSRGVLISGVLHKAFVETIEGAEAAATVAVAFSS---- 355
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 347 KQLPOST--LFRADHPFLFVIR--KDDIILFSGKVSCEP 380
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 356 ---PTSTNEEPHCNHPFLFFIRQNTKNSILFYGRFSSP 390
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 13
US-09-823-187-30
; Sequence 30, Application US/09823:187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-30

Query Match 38.4%; Score 752; DB 11; Length 390;
Best Local Similarity 41.5%; Pred. No. 1e-60;
Matches 165; Conservative 81; Mismatches 126; Indels 26; Gaps 10;
QY 1 MASLAANAECFNLPREMDNQGNGNVFFSSLSLFAALALVRLGAQDLSLQIDKLHV 60
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MNSLSEANTKFMFDLQQQRKSKEN-NIFYSPISITSALGWLGLGAKDNTAQIKKVLHF 59
```

```
QY 61 -----NTASGYGNSNSQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDY 114
Db 60 DQVTEINTGKATYHVDSEGNVHHQFQKLLTFEPNKSDDAYELKIANKLFGEKTYLFQY 119
QY 115 IECAEKLYDAKVERVDFTNHLIEDTRNNKXVNEETHGKIKNVIGEGISSAVVMVLNA 174
Db 120 LDAIKKFYQTSVESVDANAPESRKKINSWESQTNKIKNLPEGNIGSNTTLVLNA 179
QY 175 VYFKGWQSAFTKSTINCIFPKPGSGKAVAMHQRKENLSVIEDPSMKILELRNG- 233
Db 180 IYFKGWQKXFKNEKDEKFWPNKNKYKSIQMMRQYTSFHFASLEDVQAKVLEIPYKKG 239
QY 234 GINMYVLLPE--NDLSEIENKLTFOHMEWTNPRMTSKYVEVFPFKIEKNYEMKQYL 291
Db 240 DLSMIVLLPNEIDGLQLEKLTAEKLNENWTSQNMRETRVDLHLPRFKVEESYDLKDTL 299
QY 292 RALGLKDIPEKADLSGIASGRLYISRMMEKSVIEVTEGTEATAAT-----GSNIVE 346
Db 300 RTMGWVDIFN-GDADLSGWTGSLVLSGLVHKAFVTEBEGAAAAATAVVGFGSS--- 355
QY 347 KQLPOST--LFRADHPFLFVIR--KDDILFGKVCSP 380
Db 356 ---FTSNEEFHCNHPFLFPIRQNKTNLSILFYGRSSP 390

RESULT 14
US-10-094-944-12
; Sequence 12, Application US/10094944
; Publication No. US20030180275A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Clarke, Howard R
; APPLICANT: DuBose, Robert F
; APPLICANT: Wiley, Steven R
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
; FILE REFERENCE: 3223-A
; CURRENT APPLICATION NUMBER: US/10/094,944
; CURRENT FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-944-12

Query Match 38.4%; Score 752; DB 12; Length 390;
Best Local Similarity 41.5%; Pred. No. 1e-60;
Matches 165; Conservative 61; Mismatches 126; Indels 26; Gaps 10;

QY 1 MASLAAANAEFCNLFREMDNDQNGNVPFSSLSLFAALALVRLGAQDLSQIDKLHLV 60
Db 1 MNSLSEANTKMPDLFQFRKSKEN-NIFYSPISITSSALGWLLGAKDNTAQIKKVLHF 59
QY 61 -----NTASGYGNSNSQSG-LOSQKRVFSDINASHKDYDLSIVNGLFAEKVYGFHKDY 114
Db 60 DQVTEINTGKATYHVDSEGNVHHQFQKLLTFEPNKSDDAYELKIANKLFGEKTYLFQY 119
QY 115 IECAEKLYDAKVERVDFTNHLIEDTRNNKXVNEETHGKIKNVIGEGISSAVVMVLNA 174
Db 120 LDAIKKFYQTSVESVDANAPESRKKINSWESQTNKIKNLPEGNIGSNTTLVLNA 179
QY 175 VYFKGWQSAFTKSTINCIFPKPGSGKAVAMHQRKENLSVIEDPSMKILELRNG- 233
Db 180 IYFKGWQKXFKNEKDEKFWPNKNKYKSIQMMRQYTSFHFASLEDVQAKVLEIPYKKG 239
QY 234 GINMYVLLPE--NDLSEIENKLTFOHMEWTNPRMTSKYVEVFPFKIEKNYEMKQYL 291
Db 240 DLSMIVLLPNEIDGLQLEKLTAEKLNENWTSQNMRETRVDLHLPRFKVEESYDLKDTL 299
QY 292 RALGLKDIPEKADLSGIASGRLYISRMMEKSVIEVTEGTEATAAT-----GSNIVE 346
Db 300 RTMGWVDIFN-GDADLSGWTGSLVLSGLVHKAFVTEBEGAAAAATAVVGFGSS--- 355
```

```
QY 347 KQLPOST--LFRADHPFLFVIR--KDDILFGKVCSP 380
Db 356 ---FASTNEEFHCNHPFLFPIRQNKTNLSILFYGRSSP 390

RESULT 15
US-10-113-113-2
; Sequence 2, Application US/10113113
; Publication No. US20030166852A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: PROTEINASE INHIBITOR ZSERP9
; FILE REFERENCE: 00-92
; CURRENT APPLICATION NUMBER: US/10/113,113
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 60/280,678
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-113-113-2

Query Match 38.3%; Score 750.5; DB 12; Length 425;
Best Local Similarity 37.4%; Pred. No. 1.e-60;
Matches 159; Conservative 86; Mismatches 135; Indels 45; Gaps 5;

QY 1 MASLAAANAEFCNLFREMDNDQNGNVPFSSLSLFAALALVRLGAQDLSQIDKLHLV 60
Db 1 MDSLVTANTKPCFDFQEGKDRHRKNIFPSLUSALGWLRLGARSDSAHQIDEVLHP 60
QY 61 NTAS-----GYGNSNSQSGSQS-QKRVF 84
Db 61 NEFSQNESKEPDCLSKNKQKVLADSSLEQKKTTPDLDDQAGSLNNSGLVSYFGQLL 120
QY 85 SDINASHKDYDLSIVNGLFAEKVYGFHKDYIECAEKLYDAKVERVDFTNHLIEDTRNIN 144
Db 121 SKLDRIKTDYTLSTANRLYGEQEPFCQYLDGVIOFYHTTIESVDYFQKPEKSKOEINF 180
QY 145 WVENETHGKIKNVIGEGISSAVVMVLNAVYPKGWQSAFTKSETINCHFKSPKCSGKA 204
Db 181 WVECSQKIKELPSKDINAETVLVLNAVYFAKWETYPHENTVDAPFCLNANEKS 240
QY 205 VAMMQRKFNLSVIEDPSMKILELY-NGGINMYVLLPEND-----LSEIENKLTQON 257
Db 241 VKWMTQKGLYRIGFIEVKAQILEMRYTKGKLSMFWLLPSHSKONLKGLELEKRTHEK 300
QY 258 LMEWTNPRMTSKYVEVFPFKIEKNYEMKQYLRLGKLDIFDESKADLSGIASGRLY 317
Db 301 MVAMSSSENMEESVLSVSPFTLESDYDLSIQDMGITDIFDETADLUTGISPSFNL 360
QY 318 ISSMHKSVIEVTEGTEATAATGSNIVEKQLPQSTLFRADHPFLFVIR--KDDILFGS 375
Db 361 LSKLHKTFVVDENGIGQAAATGAVSERSLSWVEFNANHPFLFPIRHNKQTLIFVG 420
QY 376 KVCSP 380
Db 421 RVCSP 425

Search completed: December 12, 2003, 16:24:31
Job time : 32 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:20:11 ; Search time 17 Seconds  
(without alignments)  
1051.185 Million cell updates/sec

Title: US-09-508-997A-2

Perfect score: 1959

Sequence: 1 MASLAAANAEPCNLFREND.....FLVIRKDDILFSGKVSCTP 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1959	100.0	380	1 SPB7_HUMAN	Q75635 homo sapien
2	1505	76.8	380	1 SPB7_MOUSE	Q96955 mus musculu
3	819	41.8	392	1 SB11_HUMAN	Q96P15 homo sapien
4	753.5	39.0	405	1 SB12_HUMAN	Q96P63 homo sapien
5	752	38.4	390	1 SCC1_HUMAN	P95088 homo sapien
6	748	38.2	390	1 SCC2_HUMAN	P48594 homo sapien
7	739.5	37.7	379	1 ILEU_HORSE	P05619 equus cabal
8	722.5	36.9	379	1 ILEU_HUMAN	P30740 homo sapien
9	714.5	36.5	391	1 SB13_HUMAN	Q8U1V8 homo sapien
10	711	36.3	388	1 OVAY_CHICK	P01014 gallus gall
11	702	35.8	378	1 PT16_BOVIN	O02739 bos taurus
12	696	35.5	378	1 ILEU_FIG	P80229 sus scrofa
13	664.5	33.9	415	1 PA12_HUMAN	P05120 homo sapien
14	662	33.8	374	1 SPB8_HUMAN	P50452 homo sapien
15	658.5	33.6	397	1 SB10_HUMAN	P48595 homo sapien
16	654	33.4	385	1 OVAL_CHICK	P01012 gallus gall
17	633.5	33.4	382	1 OVAL_COTUA	P39104 coturnix co
18	649	33.1	378	1 PT16_MOUSE	Q60854 mus musculu
19	644	32.9	376	1 PT16_HUMAN	P35237 homo sapien
20	630.5	32.2	415	1 PA12_MOUSE	P12388 mus musculu
21	623	31.8	376	1 SPB9_HUMAN	P50453 homo sapien
22	622	31.8	416	1 PA12_RAT	P39524 rattus norv
23	605.5	30.9	375	1 MASP_RAT	P70564 rattus norv
24	600.5	30.7	375	1 MASP_MOUSE	P70124 mus musculu
25	592	30.2	465	1 ANT3_MOUSE	P32261 mus musculu
26	590.5	30.1	375	1 MASP_HUMAN	P36952 hmo sapien
27	567	28.9	433	1 ANT3_BOVIN	P41361 bos taurus
28	566	28.9	464	1 ANT3_HUMAN	P01008 homo sapien
29	557	28.4	465	1 ANT3_SHEEP	P32282 ovis aries
30	503.5	25.7	410	1 NEUS_RAT	Q9J182 rattus norv
31	498.5	25.4	410	1 NEUS_HUMAN	Q99574 homo sapien
32	492	25.1	405	1 SPI2_MOUSE	Q9Jk88 mus musculu
33	484.5	24.7	410	1 NEUS_MOUSE	Q35684 mus musculu

## RESULT 1

SPB7\_HUMAN

ID SPB7\_HUMAN STANDARD; PRT; 380 AA.

AC Q75635;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Megsin (TP55) (Serpin B7).

GN SERPINB7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_TaxID:9606;

[1]

SEQUENCE FROM N.A.

RA MEDLINE-97326116; PubMed-9182567;

RA Tsujimoto M., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,

RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,

RA Katayama T., Nakao M., Yamauchi K., Hashino J., Haruyama M., Miura K.,

RA Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yanaguchi N.,

RA "Purification, cDNA cloning, and characterization of a new serpin with

RT megakaryocyte maturation activity."

RL J. Biol. Chem. 272:15373-15380(1997).

[2]

SEQUENCE FROM N.A.

TISSUE-Mesangial cells;

RA MEDLINE-98376452; PubMed-9710452;

RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,

RA Okubo K., Kurokawa K.;

RA "A mesangium-predominant gene, megsin, is a new serpin upregulated in

RL IGA nephropathy."

RL J. Clin. Invest. 102:828-836(1998).

CC -!- FUNCTION: Might function as an inhibitor of Lys-specific

proteases. Might influence the maturation of megakaryocytes via

CC its action as a serpin.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.

CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; D88575; BAA31232.1; -.

EMBL; AF027866; AAC64506.1; -.

HSSP; P05619; IHLE.

Genew; HGNC:13902; SERPINB7.

MM; 603357; -.

GO; GO:0004868; F:serpin; TAS.

InterPro; IPR000215; Serpin.

Pfam; PF00079; serpin; 1.

SMART; SM00093; SERPIN; 1.

```

DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CFE05 CRC64;

Query Match 100.0%; Score 1959; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 4e-129;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAARFCFNLPREMDNQNGNVFSSLSLFAALALVRLGAGDLSQIDKLHV 60
DB 1 MASLAANAARFCFNLPREMDNQNGNVFSSLSLFAALALVRLGAGDLSQIDKLHV 60
QY 61 NTASGYGNSNSQSGQLQSKRVFSDINASHKDYDLSVNGLFAEKYGFHGYIECAEK 120
DB 61 NTASGYGNSNSQSGQLQSKRVFSDINASHKDYDLSVNGLFAEKYGFHGYIECAEK 120
QY 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKNVI GEGGISSSAVMVLNAVYFKG 180
DB 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKNVI GEGGISSSAVMVLNAVYFKG 180
QY 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKNVI GEGGISSSAVMVLNAVYFKG 180
DB 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKNVI GEGGISSSAVMVLNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSGKAVMMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVMMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
QY 181 WSAFTKSETINCHFKSPKCSGKAVMMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVMMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
QY 241 LPENDLSIEINKLTFQNLMEWTPRMTSKYVEVFPQFKIEKYNEMKQYLRAIGKDI 300
DB 241 LPENDLSIEINKLTFQNLMEWTPRMTSKYVEVFPQFKIEKYNEMKQYLRAIGKDI 300
QY 241 LPENDLSIEINKLTFQNLMEWTPRMTSKYVEVFPQFKIEKYNEMKQYLRAIGKDI 300
DB 241 LPENDLSIEINKLTFQNLMEWTPRMTSKYVEVFPQFKIEKYNEMKQYLRAIGKDI 300
QY 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2
SPH7 MOUSE
ID SPH7 MOUSE STANDARD; PRT; 380 AA.
AC Q9D695;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Megsin (Serpin 87).
GN SERPINB7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21368006; Pubmed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.:
RT "Cloning of rodent megsin revealed its up-regulation in
RT mesangio proliferative nephritis."
RL Kidney Int. 60:641-652 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STEAIN=C57BL/6J; TISSUE=SKIN;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

```

```

Browstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hornam M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordore P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoerbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
Hayashizaki Y.:
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC -!- FUNCTION: Might function as an inhibitor of Lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF105328; AAL16768.1; -.
CC EMBL; AK014524; BAE29410.1; -.
CC HSSP; P05121; 1A7C.
CC MGD; MGI:2151053; Serpinb7.
CC InterPro; IPR000215; Serpin.
CC Pfam; PF00079; serpin; 1.
CC SMART; SM00093; serpin; 1.
CC PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCE9CF4 CRC64;

Query Match 76.8%; Score 1505; DB 1; Length 380;
Best Local Similarity 73.9%; Pred. No. 1.4e-97;
Matches 281; Conservative 53; Mismatches 46; Indels 0; Gaps 0;

QY 1 MASLAANAARFCFNLPREMDNQNGNVFSSLSLFAALALVRLGAGDLSQIDKLHV 60
DB 1 MASLAANAARFCFNLPREMDNQNGNVFSSLSLFAALALVRLGAGDLSQIDKLHV 60
QY 61 NTASGYGNSNSQSGQLQSKRVFSDINASHKDYDLSVNGLFAEKYGFHGYIECAEK 120
DB 61 NTASGYGNSNSQSGQLQSKRVFSDINASHKDYDLSVNGLFAEKYGFHGYIECAEK 120
QY 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKNVI GEGGISSSAVMVLNAVYFKG 180
DB 121 LYDAKVERVDFTHLEDTRNINKWVENETHGKIKNVI GEGGISSSAVMVLNAVYFKG 180
QY 181 WSAFTKSETINCHFKSPKCSGKAVMMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVMMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
QY 181 WSAFTKSETINCHFKSPKCSGKAVMMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVMMHQRKFNLSVIEDPSMKILELYNGGINMYVL 240
QY 241 LPENDLSIEINKLTFQNLMEWTPRMTSKYVEVFPQFKIEKYNEMKQYLRAIGKDI 300
DB 241 LPENDLSIEINKLTFQNLMEWTPRMTSKYVEVFPQFKIEKYNEMKQYLRAIGKDI 300
QY 241 LPENDLSIEINKLTFQNLMEWTPRMTSKYVEVFPQFKIEKYNEMKQYLRAIGKDI 300
DB 241 LPENDLSIEINKLTFQNLMEWTPRMTSKYVEVFPQFKIEKYNEMKQYLRAIGKDI 300
QY 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
DB 301 DESKADLSGIASGRILYISRMHKSYLEVTEEGTATATGNSNIVEKQLPQSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 3
SB11 HUMAN
ID SB11 HUMAN STANDARD; PRT; 392 AA.
AC Q96P15; Q96P13; Q96P14;

```

```

RESULT 4
SB12_HUMAN
ID   SB12_HUMAN          STANDARD;          PRT;   405 AA.
AC   Q9P63;
DT   28-FEB-2003 (Rel. 41, Created)
DI   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Serpin B12.
GN   SERPINB12.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI TaxID=9606;
RN   [1]
RX   SEQUENCE FROM N.A., AND CHARACTERIZATION.
EX   MEDLINE=21623618; PubMed=11604408;
RA   Askew Y.S., Pak S.C., Luke C.J., Askew D.J., Cataltepe S., Mills D.R.,
RA   Kato H., Lehoczy J., Dewar K., Birren B., Silverman G.A.;
RT   "SERPINB12 is a novel member of the human ov-serpin family that is
RT   widely expressed and inhibits trypsin-like serine proteinases.";
RL   J. Biol. Chem. 276:49320-49330(2001).
CC   -!- FUNCTION: Inhibits trypsin and plasmin, but not thrombin,
CC   coagulation factor Xa, or urokinase-type plasminogen activator.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in many tissues, including brain,
CC   bone marrow, lymph node, heart, lung, liver, pancreas, testis,
CC   ovary, and intestine.
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
EMBL; AF411191; AA05571.1; -.
Genew; HGNC:14220; SERPINB12.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0019899; F:enzyme binding activity; NAS.
DR   GO; GO:0004860; F:serpin; IMP.
DR   GO; GO:0030304; F:trypsin inhibitor activity; IMP.
DR   GO; GO:0042177; P:negative regulation of protein catabolism; NAS.
DR   InterPro; IPR000215; Serpin.
DR   Pfam; PF00079; serpin; 1.
DR   SMART; SM00093; SERPIN; 1.
DR   PROSITE; PS00284; SERPIN; 1.
KW   Serpin; Serine protease inhibitor.
FT   ACT_SITE 370..371      REACTIVE BOND (BY SIMILARITY).
SQ   SEQUENCE 405 AA; 46276 MW; PFE12DC9B7F3DEA CRC64;

Query Match          39.0%; Score 763.5; DB 1; Length 405;
Best Local Similarity 39.3%; Pred No. 5,4e-46;
Matches 159; Conservative 87; Mismatches 134; Indels 25; Gaps 5;

QY 1 MASLAANAFFCNLFREMDNMGNGVFFSLSLFAALAYRLGAQDLSLQIDKLHLV 60
Db 1 MDSLVTANTKCFDLFQEIIGDKDRHKNIFFPFLSLSAALGMVRLGARDSAHOIDEVLHF 60
QY 61 NTAS-----GYGNSNSNSQGLQS-QLRVPSDINASHKDYDLISVNLGFA 104
Db 61 NFPSQNSKEPDPCLKSNKQKASGLANESGLVSCYFGLLSKLDRIKTDYTLSTANRLYG 120
QY 105 EKTVGFHKDYIECAEKLIDAKVERVDFTNHLEDPRRNKNWVENETHGKINWIGGGIS 164
Db 121 EQEFPFICQELVDGVTFPHITLESVDQFNPKESQELNFWVECSQSKLEIFPSKDALN 180
QY 165 SGAVMLVNAVYPKGWQSAFTKSETINCHFKSPKCSKGKAVMMHQERKNVLVIEDPSM 224
Db 181 AETVLVLNAVYPKAWETYPDENTVDAPFCLNANENKSNVMTQKGLYRIQFIEEVKA 240

```





DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Squamous cell carcinoma antigen 2 (SCCA-2) (leupin).  
GN SERPINB4 OR SCCA2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95241462; PubMed=7724531;  
RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.,  
RA Treter S.D., Hui S.M., Silverman G.A.;  
RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem  
RT duplication of the human squamous cell carcinoma antigen gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96013987; PubMed=7599435;  
RA Barnes R.C., Worrall D.M.;  
RA "Identification of a novel human serpin gene; cloning sequencing and  
RT expression of leupin.";  
RL FEBS Lett. 373:61-65(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21167379; PubMed=11267667;  
RA Hamada K., Shinomiya H., Asano Y., Kihara T., Iwanoto M., Hanakawa Y.,  
RA Hachimoto K., Hirose S., Kyo S., Ito M.;  
RT "Molecular cloning of human squamous cell carcinoma antigen 1 gene and  
RT characterization of its promoter.";  
RL Biochim. Biophys. Acta 1518:124-131(2001).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins B., Wagner L., Shenmen C.M., Schaller G.D.,  
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
RA Brownstein M.J., Ueffing T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST  
CC IMMUNE RESPONSE AGAINST TUMOR CELLS.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- TISSUE SPECIFICITY: Squamous cells.  
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U19576; AAA92602.1; -  
CC EMBL; U19570; AAA92602.1; JOINED.  
CC EMBL; U19571; AAA92602.1; JOINED.  
CC EMBL; U19572; AAA92602.1; JOINED.  
CC EMBL; U19574; AAA92602.1; JOINED.

DR EMBL; U19575; AAA92602.1; JOINED.  
DR EMBL; U19557; AAA97553.1; -  
DR EMBL; X89015; CAA61420.1; -  
DR EMBL; AB035089; BAB21525.1; -  
DR EMBL; BC017401; AAH17401.1; -  
DR HSSP; P01008; IATH.  
DR Genew; HGNC:10570; SERPINB4.  
DR MIM; 600518; -  
DR GO; GO:0005737; C:cytoplasm; NAS.  
DR GO; GO:0004868; F:serpin; NAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR GO; GO:0030162; P:regulation of proteolysis and peptidolysis; NAS.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin; 1.  
DR SMART; SM00093; SERPIN; 1.  
DR PROSITE; PS00284; SERPIN; 1.  
KW Serpin; serine protease inhibitor.  
FT ACT\_SITE 354..355  
FT REACTIVE\_BOND  
SQ SEQUENCE 390 AA; 44854 MW; 04E213CD892587D5 CRC64;  
Query Match 38.2%; Score 748; DB 1; Length 390;  
Best Local Similarity 40.8%; Pred. No. 6.2e-45;  
Matches 160; Conservative 84; Mismatches 134; Indels 14; Gaps 8;  
QY 1 MASLAANAARFCVLFREMDNCGNGNVFSSLSLFAALVRLGAQDLSQIDKLHV 60  
DB 1 MNSLSEANTKFMFLFQPRKSKEN-NIFYSPISITSALGMVLLGAKNTAQQISKVLHF 59  
QY 61 ----NTASGCGNSSNSQSG--LQSOLKRVFSDINASHKDYDLSIVNGLFAEKVVGFKDY 114  
DB 60 DQVTNTEKATYHVDKSGNVHGFQKLLTEFNKSTDAYELKIANKLFGEKYQFLQY 119  
QY 115 IECAEKLYDAKVERVDFTHLEDRNINKWVENETHGKIKVIGEGGISSAWMYLVNA 174  
DB 120 LDAIKFKQTSTVSTDFANAPESKKINSWESQTNKIKLFPDGTIGNDITLVNA 179  
QY 175 VYFKGHWOSATKSETINCHFSKPSCKSAVAMHGHQERFNLSVIEDPSMKILELYNG- 233  
DB 180 IYFKGOWNKPKKNTKEEFKFWNPNTYKSVQMMRQYNSFNALLEDVQAKVLEIPYK 239  
QY 234 GINNYVLLPE--NDLSIENKLTFLQNLMEWTNPRMTSKYVVEFPPOFKIKENYMKYL 291  
DB 240 DLSKIVLLENGIDGQKLEKLTAEKMEWTSIQNRETCDVLIHLPKMEESYDKDL 299  
QY 292 RALGKQIDFDSKADSGASGRIYISRMHKSXYLEVTEGTEATAATGNTIVERKQLFQ 351  
DB 300 RTGMVNIFFN-GDADLSGMTWGHLSVSKLHKAFVETEGVEAAAAATAVWVVELSSPS 358  
QY 352 ST-LFEADHPFLFVIR--KDDILFSGKVSCLP 380  
DB 359 TNEEPCNHPFLFFIRQNTKNSILFYGRFSP 390  
RESULT 7  
ID ILEU\_HORSE  
AC P05619;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Leukocyte elastase inhibitor (LEI).  
GN SERPINB1 OR ELANH2.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93319507; PubMed=7687128;  
RA Kordula T., Dubin A., Schooltink H., Koj A., Heinrich P.C.,  
RA Rose-John S.;  
RT "Molecular cloning and expression of an intracellular serpin: an  
RT elastase inhibitor from horse leucocytes.";

RL Biochem. J. 293:187-193(1993).  
 RN [2].  
 RP SEQUENCE.  
 RC TISSUE=Leukocyte; PubMed=1551869;  
 RX MEDLINE=92202200; PubMed=1551869;  
 RA Dubin A., Travis J., Englund J.J., Potempa J.;  
 RT "Equine leukocyte elastase inhibitor. Primary structure and  
 identification as a thymosin-binding protein.";  
 RL J. Biol. Chem. 267:6576-6583(1992).  
 RN [3].  
 RP SEQUENCE OF 343-362.  
 RX MEDLINE=89213423; PubMed=3366785;  
 RA Potempa J., Dubin A., Watorek W., Travis J.;  
 RT "An elastase inhibitor from equine leukocyte cytosol belongs to the  
 serpin superfamily. Further characterization and amino acid sequence  
 of the reactive center.";  
 RL J. Biol. Chem. 263:7364-7369(1988).  
 RN [4].  
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).  
 RX MEDLINE=92389327; PubMed=1518052;  
 RA Baumann U., Bode W., Huber R., Travis J., Potempa J.;  
 RT "Crystal structure of cleaved equine leukocyte elastase inhibitor  
 determined at 1.95-A resolution.";  
 RL J. Mol. Biol. 226:1207-1218(1992).  
 CC -1- FUNCTION: THIS INHIBITOR IS THOUGHT TO BE INVOLVED IN THE CONTROL  
 OF INTRACELLULAR PROTEIN TURNOVER. BINDS TO THYMOSIN BETA-4.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M91161; AAA97513.1; -.  
 DR PIR; A28060; A28060.  
 DR PIR; A2421; A2421.  
 DR PDB; 1HLE; 31-JAN-94.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; serpin; 1.  
 DR SMART; SM00093; SERPIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor; 3D-structure.  
 FT MOD\_RES 1 1  
 FT ACT\_SITE 344 345  
 FT 1 BLOCKED (WITH AN UNKNOWN CHEMICAL GROUP).  
 FT 2 REACTIVE SOND.  
 FT 344 345  
 FT 2 22  
 FT STRAND 28 30  
 FT HELIX 32 44  
 FT TURN 45 45  
 FT HELIX 48 58  
 FT TURN 59 59  
 FT HELIX 60 62  
 FT TURN 64 65  
 FT HELIX 66 77  
 FT TURN 78 78  
 FT STRAND 85 95  
 FT TURN 96 97  
 FT HELIX 102 112  
 FT STRAND 115 119  
 FT TURN 121 123  
 FT HELIX 125 139  
 FT TURN 140 142  
 FT TURN 150 151  
 FT TURN 155 156  
 FT STRAND 157 171  
 FT HELIX 177 179  
 FT STRAND 181 186  
 FT STRAND 192 209  
 FT TURN 210 212  
 FT HELIX 210 212  
 FT TURN 213 213

FT STRAND 214 221  
 FT TURN 222 223  
 FT STRAND 226 233  
 FT HELIX 244 248  
 FT TURN 249 249  
 FT HELIX 252 259  
 FT HELIX 261 263  
 FT STRAND 265 274  
 FT STRAND 276 283  
 FT HELIX 285 291  
 FT TURN 292 292  
 FT HELIX 295 297  
 FT TURN 299 301  
 FT HELIX 305 308  
 FT STRAND 313 326  
 FT STRAND 330 343  
 FT STRAND 350 353  
 FT STRAND 358 364  
 FT TURN 365 368  
 FT STRAND 369 376  
 SQ SEQUENCE 379 AA; 43006 MW; 72DED1999371427A CRC64;  
 Query Match 37.7%; Score 739.5; DB 1; Length 379;  
 Best Local Similarity 38.0%; Pred. No. 2.3e-44;  
 Matches 146; Conservative 85; Mismatches 137; Indels 19; Gaps 4;  
 QY 1 MASLAANAEEFCNLFREMDNQNGNVPFSSLSLFAALVRLGAODDSLSQIDKLHV 60  
 Db 1 MEQLSTANTHFAVDFALNESDPTGNIFISPLSISSALAMIFLGRGNTAAQVSKALYF 60  
 QY 61 NTASGYGNSNSQSGCQSQKLVPSDINASHKQYDLSIVNGLFAEKVYGFHKDYIECAEK 120  
 Db 61 DTV-----BDHSPQSLNADINKPGAPYIIKLANLYGKTYNFIADFLASTOK 110  
 QY 121 LYDAKVERVDFTNLETRNINKWVENETHGKIKNVI GEGSISSAMVVLVNAVYFKGK 180  
 Db 111 MYGAEIASVDQQAPEDARKEINENWKGTEGKIPELLVKGWVDNMTKLVNAVYFKGN 170  
 QY 181 WQSAFTKSETINCHFKSPKCSGKAVAMQHQRKPNLSVIEDPSMKILELYNG-GINMYV 239  
 Db 171 WQEKPMKEATRDAPFLNKDQTKTVKMYQKKFPYNYIEDLKCVLELFPYQKLSMII 230  
 QY 240 LIPEN-----DLSEIENKLTQNLMEWTNPRMTSKYVEVFPQFKIKNYEMKQYLRA 293  
 Db 231 LLPDDIEDSTGLSEIKQTLLEKREWKPNENLYLAENVNHLPRKLESYDLTSLHAR 290  
 QY 294 LGLKDIFDESKADLSGASGRGIYISRMHKSXYIEVTGTEATATAGTSNIVKQLPQST 353  
 Db 291 LGVQDLFNRGKADLSGMSGARDLFVSKIIHKSPVDLNEGTEAAAAATAGTITLMLMPZE 350  
 QY 354 LFRADHPFLFVIRKDDI--ILFSGKVSQCP 380  
 Db 351 NFNADHPFIFFIRHNPNSANILFLGRFSSP 379  
 RESULT 8  
 ID ILEU HUMAN STANDARD; PRT; 379 AA.  
 AC P30740;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Leukocyte elastase inhibitor (LEI) (Monocyte/neutrophil elastase  
 inhibitor) (M/NEI) (H1).  
 GN SERPINB1 OR ELANH2 OR PI2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92302296; PubMed=1376927;  
 RA Remold-O'Donnell B., Chin J., Alberts M.;



RI Biochim. Biophys. Acta 1492:441-446(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Placenta;  
 RX MEDLINE=21213570; PubMed=11313015;  
 RA Acts H.F., Weiss T., Scheuring S., Scott F.J., Irving J.A., Michel G.,  
 RA Bards P.I., Ruzicka T.;  
 RT "Sequence, organization, chromosomal localization and alternative  
 RT splicing of the human serine protease inhibitor gene hurpin (P113),  
 RT which is up-regulated in psoriasis.";  
 RL DNA Cell Biol. 20:123-131(2001).  
 CC -|- FUNCTION: May play a role in the proliferation or differentiation  
 CC of keratinocytes.  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -|- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9UIV8-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9UIV8-2; Sequence=VSP\_006058;  
 CC TISSUE SPECIFICITY: SKIN-SPECIFIC.  
 CC -|- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AJ001696; CAA04935.2; -;  
 CC EMBL; AJ001697; CAA04936.2; -;  
 CC EMBL; AJ001698; CAA04937.1; -;  
 CC EMBL; AF169949; AAD55765.1; -;  
 CC EMBL; AF216854; AAF27879.1; -;  
 CC EMBL; AJ278717; CAC03569.1; -;  
 CC PIR; JC7118; JC7118.  
 CC HSP; P05120; 1BY7.  
 CC Genew; HGNC:8944; SERPINB13.  
 CC MIM; 604445; -;  
 CC GO; GO:0004868; F:serpin; NAS.  
 CC GO; GO:0030162; P:regulation of proteolysis and peptidolysis; NAS.  
 CC GO; GO:0009411; P:response to UV; TAS.  
 CC InterPro; IPR000240; Maspin.  
 CC InterPro; IPR000215; Serpin.  
 CC Pfam; PF00079; serpin; 1.  
 CC PRINTS; PR00676; MASPIN.  
 CC SMART; SM00093; SERPIN; 1.  
 CC PROSITE; PS00284; SERPIN; 1.  
 CC Serpin; Serine protease inhibitor; Alternative splicing.  
 KW ACT SITE 356 357 REACTIVE BOND (BY SIMILARITY).  
 FT VARSPLIC 206 257 Missing (in isoform 2).  
 FT FTID=VSP\_006058.  
 FT  
 FT CONFLICT 8 8 S -> N (IN REF. 4).  
 FT CONFLICT 75 75 MISSING (IN REF. 1; CAA04937).  
 FT CONFLICT 293 293 G -> S (IN REF. 2).  
 FT CONFLICT 297 297 E -> Q (IN REF. 1; CAA04937).  
 SQ SEQUENCE 391 AA; 44276 MW; 2CA88558D4BC2B09 CRC64;  
 Query Match 36.5%; Score 714.5; DB 1; Length 391;  
 Best Local Similarity 38.0%; Pred. No. 1.3e-42;  
 Matches 152; Conservative 77; Mismatches 142; Indels 29; Gaps 7;  
 QY 1 MASIAAANAFCNLFREMDNDGNGNFFSLSLFAALALVRLGAQDCSLSDIKLLHV 60  
 Db 1 YDSUGAVSLGFLDFXEL-KKINDGNIFPSPVGLITAGVLLGTGATASQLEBVFH- 58  
 QY 61 NTASGVGNSNSQSGLOSLKRVFSDINASHK-----DYDLSLVNGLPAEK 106  
 Db 59 -----SEKETSSRIKAEKEVIENTEAVHQCPQKFLTEISKLTNDYELNITNRLFGK 112  
 QY 107 VYGHKDYECARLYDAKVERVDFTNLEDTREINKNWENETHGKIKNVIGEGGISS 166

Db 113 TVLFLQKLYDVEKYYHASLEPVDVNADESRKKNWSKTKNKLKDLFPDGSISS 172  
 QY 167 AMVLVNAVYFKGQWOSAFKSETINCHFKSPKCSGKAVAMHQERKFNLSVIRDPNMI 226  
 Db 173 TKLVLVNMVYFKQWDRFCKKENTKEEFWANKSTSKSVQMTQSHSFSTFLEDLQAKI 232  
 QY 227 LELRY-NGGNNMYVLLPNDLSEIE--NKLTFQNLMENTNPRMTSKVVEFFPQPKIE 282  
 Db 233 LGIPYXNDLSMFVLLP-NDIDGLEKIIDKIPKLVWNTSPGHMEERKVNHLRFRFEVE 291  
 QY 283 KNYEMKQVLRALGLKIDPESKADLSGAGRLYISRMHKSIVIEVTEGTEATAATGS 342  
 Db 292 DGDLEAVLAMNGDAPSEKADYSGMSSGSLYAKFLHSSFVATBEGTEAAAAGI 351  
 QY 343 NIVEKQLPQSTLFRADHPFLFVIR--KDDILLFSKVKSCP 380  
 Db 352 GFTVTSAPGHENVCNHPFLFFIRHNSNLSILFFGRFSSP 391  
 RESULT 10  
 OVAY CHICK  
 ID OVAY CHICK STANDARD; PRT; 388 AA.  
 AC P01014;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Gene Y protein (Ovalbumin-related).  
 GN Y.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=83014329; PubMed=7122240;  
 RA Heilig R., Muraskowsky R., Kloepper C., Mandel J.L.;  
 RT "The ovalbumin gene family: complete sequence and structure of the Y  
 RT gene.";  
 RL Nucleic Acids Res. 10:4363-4382(1982).  
 CC -|- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; J00922; AAA68882.1; -;  
 CC PIR; A01244; DYCH.  
 CC HSP; P01012; LOVA.  
 CC InterPro; IPR000215; Serpin.  
 CC Pfam; PF00079; serpin; 1.  
 CC SMART; SM00093; SERPIN; 1.  
 CC PROSITE; PS00284; SERPIN; 1.  
 KW Serpin.  
 FT ACT SITE 353 354 REACTIVE BOND HOMOLOG.  
 FT CARBOHYD 95 95 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 215 215 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 312 312 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 388 AA; 43772 MW; 2AF1BF715D7F461 CRC64;  
 Query Match 36.3%; Score 711; DB 1; Length 388;  
 Best Local Similarity 39.6%; Pred. No. 2.3e-42;  
 Matches 155; Conservative 75; Mismatches 147; Indels 14; Gaps 7;  
 QY 1 MASIAAANAFCNLFREMDNDGNGNFFSLSLFAALALVRLGAQDCSLSDIKLLHV 60  
 Db 1 MDSISVNAKFCDFVFMKVVHVNILYCPLSILTALAMVYLARGNTESQMKVLHF 60

QY	61	NTASGYGNSNSQSQ----	LOSQKRVFSDINASHKDYDLSIVNGLFASKVYGFHKDYIE	116
DB	61	DSITGAGTTDQSGSSEYVHNL	FKELLSEITRPNATYSLEADKLVDKTFSLPEYLS	120
QY	117	CAEKLYDAKVERVDFTNHLED	TRRNINKNWVENHGTGKINVI	176
DB	121	CARKFYTGVEVNFYTAABEA	QJLNSWKEKTEGTQKDLLVSSIDFGTMTWFIINTY	180
QY	177	FKGKQSAFTSETINCHFKSP	KSGKAVAMQHOERKKNLSVIEDPSMKILELRY	235
DB	181	FXGIWKIAPNTRETPSMTK	ESFPVQCMNSFNVAITPAEKMLLELPASGDL	240
QY	236	NMYVLILPE--NDLSEIN	KLTFQNLMEWTNPRMTSKYEVFPQPKIEKXEMKQYLR	293
DB	241	SMVLVLPDEVSGLERIEK	INTFDKLREWTSTNMAKSKMKVYLPKMKIEKNTLSILMA	300
QY	294	LGKLKIDFSDKADLSGIAS	GLRGLYISRMHKSIVIEFTGTEATATGS--NIVEKQLPQ	351
DB	301	LGMTDLFRS--ANLTG	ISVDNLMTSDAVGFVMEVNBEGTEATGTGTAIGNI--KHSLS	357
QY	352	STILFRADHPFLFVIRK	QD--DIIIFSGKVSQCP	380
DB	358	LSEFRADHPFLFIRNP	TNALIFGRYWS? 366	
RESULT 11				
ID	PT16 BOVIN	STANDARD;	PR?	378 AA.
AC	002739;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Serine proteinase inhibitor B-43.			
OS	Bos taurus (Bovine)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Retheraia; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI TaxID:9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RT	TISSUE=Brain;			
RA	MEDLINE=97165893; PubMed=9013786;			
RX	Nakaya N., Nishibori M., Kawabata M., Saeiki K.;			
RT	"Cloning of a serine proteinase inhibitor from bovine brain:			
RT	expression in the brain and characterization of its target			
RT	proteinases.";			
RL	Brain Res. Mol. Brain Res. 42:293-300(1996).			
CC	-!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES			
CC	PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.			
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-!- TISSUE SPECIFICITY: BRAIN.			
CC	-!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announcement">http://www.isb-sib.ch/announcement</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
DR	EMBL; D55670; BAA19875.1; -			
DR	HSSP; P05120; 1B77.			
DR	InterPro; IPR000215; Serpin.			
DR	Pfam; PF00079; serpin.1.			
DR	SMART; SM00053; SERPIN.1.			
DR	PROSITE; PS00284; SERPIN.1.			
KW	Serpin; Serine protease inhibitor.			
FT	ACT_SITE 343 344			
FT	SEQUENCE 378 AA; 42560 MW; 664F99CCFC263A CRC64;			
Query Match				35.8%
				Score 702; DB 1; Length 378;

```
QY 1 MASLAANAABFCNLPREMDDNGNNGVFFSLSLFAALALVRLGAQDDSLSDIKLHV 60
DB 1 MEQLSAANTFALDIFRALNESPNAGNIFISPGISSALAMILLGTRGNTEAQQSKALHF 60
QY 61 NTASGCGNSNSQSLQSLKRVFSDINASHKDYDLSIVNGLFAEKVYGHKQYIECAEK 120
DB 61 DTV-----XDIHSRPOSINADINCKGASVILKLANLFGKTHFDFPELASTOK 110
QY 121 LYDAKVERVDFNHLIEDTRNINKWVENETHGKIKNVIGRGITSSAVWVIVAVFFKKG 180
DB 111 TYGALASVDFPLASEERAKALNEWKEQTEBKIPELLASGVDSATKLVLNAIYFKGS 170
QY 181 WQSFTKSTTINCHFKSPKSGKAVAMMEQERKNLSVIEDPSMKILRLYNG-GINCVY 239
DB 171 WOEFMTATEKDAPFLNKKDSKTVKQMYQKKKPPFGYIKELKCRVLELPYQGRDLSMWI 230
QY 240 LIPEN-----DLSRIENKLTQNLMEWTPRMTSKVVEFPFQPKIEKVENMKQYLRA 293
DB 231 LLPDSIEDSTGLRKIQHUTLEKINWTKPDNLELVNHLPRFLRESYDNLNAPLAR 290
QY 294 LGLKDIFDESADLSGASGRLYISRMVHKSYIEVTEEGTEATAATGNSNIVEKQLPQST 353
DB 291 LGVQDLFG-SRADLTCGSEARDLFISKVGHKAFVNEEGTEAAATAGIAVAFMLMPEE 349
QY 354 LFRADHPFLVTKD--DILFSGKVSCP 350
DB 350 DFIADHPFIPIRHNPSSNIFLGRSSP 378

RESULT 13
PAI2 HUMAN
ID AC P05120; Q96E36; STANDARD; PRT; 415 AA.
DT 13-AUG-1987 (Rel. 05, Created)
DI 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen activator inhibitor-2 precursor (PAI-2) (Placental
DE plasminogen activator inhibitor) (Monocyte Arg-serpin) (Urokinase
DE inhibitor).
GN SRPINE2 OR PAI2 OR PLANH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=87137674; PubMed=3029122;
RA Ye R.D., Wu T.-Z., Sadler J.E.;
RT "cDNA cloning and expression in Escherichia coli of a plasminogen
RT activator inhibitor from human placenta.";
RL J. Biol. Chem. 262:3718-3725 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88142852; PubMed=3325928;
RA Schleuning W.-D., Medcalf R.L., Hession C., Rothenbuhler R.,
RA Shaw A., Kruthof E.K.O.;
RT "Plasminogen activator inhibitor 2: regulation of gene transcription
RT during phorbol ester-mediated differentiation of U-937 human
RT histiocytic lymphoma cells.";
RL Mol. Cell. Biol. 7:4564-4567 (1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174589; PubMed=2494165;
RA Ye R.D., Aherm S.M., le Beau M.M., Lebo R.V., Sadler J.E.;
RT "Structure of the gene for human plasminogen activator inhibitor-2.
RT The nearest mammalian homologue of chicken ovalbumin.";
RL J. Biol. Chem. 264:5495-5502 (1989).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Monocytes;
RX MEDLINE=87252928; PubMed=3496414;
RA Webb A.C., Collins K.L., Snyder S.F., Alexander S.J.,
RA ERRATUM.
```

RA Rosenwasser L.J., Eddy R.L., Shows T.B., Auron P.E.;  
RT "Human monocyte Arg-Serpin cDNA. Sequence, chromosomal assignment,  
RT and homology to plasminogen activator-inhibitor."  
J. Exp. Med. 166:77-94 (1987).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Monocytes;  
RX MEDLINE=88125032; PubMed=3257578;  
RA Antalis T.N., Clark M.A., Barnes T., Lehrbach P.R., Devine P.L.,  
RA Chezov G., Goss N.H., Stephens R.W., Tolstoshev P.;  
RT "Cloning and expression of a cDNA coding for a human monocyte-derived  
RT plasminogen activator inhibitor."  
Proc. Natl. Acad. Sci. U.S.A. 85:985-989 (1988).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90152678; PubMed=2303256;  
RA Samia J.A., Alexander S.J., Horton K.W., Auron P.E., Byers M.G.,  
RA Shows T.B. Jr., Webb A.C.;  
RT "Chromosomal organization and localization of the human urokinase  
RT inhibitor gene: perfect structural conservation with ovalbumin."  
J. Exp. Med. 171:159-167 (1990).  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=1247732;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Brance C.,  
RA Rana S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bonfield G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences."  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [8]  
RP SEQUENCE OF 12-17; 103-108 AND 314-321.  
RX MEDLINE=93162043; PubMed=1286667;  
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,  
RA Vandekerckhove J.;  
RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
RT protein database of normal human epidermal keratinocytes."  
J. Electrophoresis 13:960-969 (1992).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99148112; PubMed=10368272;  
RA Harrop S.J., Jankova L., Coles M., Jardine D., Whittaker J.S.,  
RA Gould A.R., Meister A., King G.C., Mabbott B.C., Curmi P.M.G.;  
RT "The crystal structure of plasminogen activator inhibitor 2 at 2.0-A  
RT resolution: implications for serpin function."  
Structure 7:43-54 (1999).  
RN [10]  
RP VARIANTS ASP-120; HUS-229; LYS-404 AND CYS-413.  
RX MEDLINE=99318093; PubMed=10391209;  
RA Cargill M., Aitken D., Ireland J., Sklar P., Ardlie K., Patil N.,  
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,  
RA Lander E.S.;  
RT "Characterization of single-nucleotide polymorphisms in coding regions  
RT of human genes."  
Nat. Genet. 22:231-238 (1999).  
RN [11]  
RP ERRATUM.

RA Cargill M., Altbuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nimesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 CC -1- FUNCTION: PAI-2 INHIBITS UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 CC THE MONOCYTE DERIVED PAI-2 IS DISTINCT FROM THE ENDOTHELIAL  
 CC CELL-DERIVED PAI-1.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC OR EXTRACELLULAR.  
 CC -1- PTM: THE SIGNAL SEQUENCE IS NOT CLEAVED.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; J02685; AAA36413.1; -;  
 CC EMBL; M18082; AAA60006.1; -;  
 CC EMBL; Y00630; CAA68666.1; -;  
 CC EMBL; M24657; AAA60348.1; -;  
 CC EMBL; M24651; AAA60348.1; JOINED.  
 CC EMBL; M24652; AAA60348.1; JOINED.  
 CC EMBL; M24653; AAA60348.1; JOINED.  
 CC EMBL; M24654; AAA60348.1; JOINED.  
 CC EMBL; M24655; AAA60348.1; JOINED.  
 CC EMBL; M24656; AAA60348.1; JOINED.  
 CC EMBL; J03603; AAA60004.1; -;  
 CC EMBL; M31551; AAA36797.1; -;  
 CC EMBL; M31547; AAA36797.1; JOINED.  
 CC EMBL; M31548; AAA36797.1; JOINED.  
 CC EMBL; M31549; AAA36797.1; JOINED.  
 CC EMBL; M31550; AAA36797.1; JOINED.  
 CC EMBL; B0012609; AAI12609.1; -;  
 CC EMBL; A21238; CAA01535.1; -;  
 CC EMBL; A21254; CAA01539.1; -;  
 CC PIR; A32853; A32853.  
 CC PDB; 1BY7; 24-OCT-99.  
 CC PDB; 1JRR; 18-DEC-02.  
 CC Aarhus/Shent-2DPAGE; 6314; IEP.  
 CC Genew; HGNC:8584; SERPINE2.  
 CC MIM; 173390; -;  
 CC GO; GO:0006916; P:anti-apoptosis; TAS.  
 CC InterPro; IPR000215; Serpin.  
 CC Pfam; PF00079; serpin; 1.  
 CC SMART; SM00093; SERPIN; 1.  
 CC PROSITE; PS00284; SERPIN; 1.  
 KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;  
 KW Glycoprotein; Signal; 3D-structure; Polymorphism.  
 FT SIGNAL 1 ? NOT CLEAVED.  
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT ACT\_SITE 380 381 REACTIVE BOND.  
 FT VARIANT 120 120 N -> D (IN dbSNP:6098).  
 FT VARIANT 229 229 R -> H (IN dbSNP:6100).  
 FT VARIANT 404 404 N -> K (IN dbSNP:6103).  
 FT VARIANT 413 413 S -> C (IN dbSNP:6104).  
 FT CONFLICT 170 170 N -> Y (IN REF. 7).  
 FT HELIX 4 22  
 FT STRAND 28 30  
 FT HELIX 32 44  
 FT TURN 45 45  
 FT HELIX 48 57  
 FT TURN 58 63  
 FT HELIX 100 102

FT HELIX 103 114  
 FT STRAND 123 132  
 FT TURN 133 134  
 FT STRAND 137 137  
 FT HELIX 139 149  
 FT STRAND 154 156  
 FT HELIX 158 176  
 FT TURN 177 179  
 FT TURN 182 183  
 FT TURN 187 188  
 FT TURN 192 193  
 FT STRAND 196 204

Query Match 33.9%; Score 664.5; DB 1; Length 415;  
 Best Local Similarity 34.7%; Pred. No. 4.3e-39;  
 Matches 144; Conservative 83; Mismatches 153; Indels 35; Gaps 5;

QY 1 MASLAAANAEFCFLFREDDNQGNGNVFSSLSLFAALALVRLGAQDLSLQDKLHV 60  
 DB 1 MEDLCVANTLFLALNLFKHLAKASPTQNLFLSPWSISSTWAMYMGSRGSTDQMAKVLQF 60

QY 61 N-----TASGY-----CN-----SSNSQSGLOSQKRVFSDINASHKD 93  
 DB 61 NEVCANAVTPTPENFTSCGFMOQIQKQSYDAILQAQADKIHSSFSLSAINASTGN 120

QY 94 YLSIVNGLFARVYGFHKYIECAEKLYDAKVERVDFTNHLIEDPRNINKWVENETHGK 153  
 DB 121 YLLESVNLFGKSGASPREYIRLQKYISSEPCQAVDFLECAEARKKINSVKTQTKGK 180

QY 154 IKNVIGEGISSAWVLNNAVYFKWQSAFTKSTINCHPKSPKSGKAVAMHEQERK 213  
 DB 181 IPNLFPSSVDGTRWLVNAVYFKGWKTFEKGNGLYPFVNSAQRTFVQWYLRK 240

QY 214 FNLSTVEDPSMKLELYNGINMYVLPE-----NDLSEIKLTFONLMEMWNPFRM 267  
 DB 241 LNTGYIEDLKAQILELPYAGDYSMLPLLPDLADVDSTGLSEIYDYKLNKTSKDM 300

QY 268 TSKYVEVFPQFKLEKNYKQYLRALGLKIDFDESKADLSGAGGLXISRMHKSXY 327  
 DB 301 AEDEVVVPQKLEHYYELSLRSMGMDAENKGRANFSGMSERNDLFLSEVFFQAMV 360

QY 328 EVTEECTATATGNSNIVEKQLPQSTLFRADHPFLFVI--RKDDILLFSGKVSQP 380  
 DB 361 DVNEEGTAAAGTGGVMTGRTGHGGPQFVADHPFLFMHKITNCILUFFGRFSPP 415

RESULT 14  
 SPB8 HUMAN  
 ID SPB8 HUMAN STANDARD; PRT; 374 AA.  
 AC P50452;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytoplasmic antiprotease 2 (CAP2) (CAP-2) (Protease inhibitor 8)  
 DE (Serpin B8).  
 GN SERPINB8 OR PI8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=96102039; PubMed=8530382;  
 RA Sprecher C.A., Morgenstern K.A., Mathewes S., Dahlen J.R.,  
 RA Schrader S.K., Foster D.C., Kiesel W.;  
 RT "Molecular cloning, expression, and partial characterization of two  
 RT novel members of the ovalbumin family of serine proteinase  
 RT inhibitors.";  
 RL J. Biol. Chem. 270:29854-29861(1995).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: L40377; AAC41939.1; --  
 CC PIR: A59273; A59273.  
 CC HSP: P05120; IRY7.  
 CC Genew: HGNC:8952; SERPINB8.  
 CC MIM: 601697; --  
 CC GO: GO:0005829; Cytosol; TAS.  
 CC GO: GO:0005515; F:protein binding activity; TAS.  
 CC GO: GO:0004868; F:serpin; TAS.  
 CC InterPro: IPR000215; Serpin.  
 CC Pfam: PF00079; serpin; 1.  
 CC SMART: SM00093; SERPIN; 1.  
 CC PROSITE: PS00284; SERPIN; 1.  
 CC Serpin; Serine protease inhibitor.  
 CC ACT\_SITE 339 340 REACTIVE BOND (BY SIMILARITY).  
 CC SEQUENCE 374 AA; 42786 MW; E855A033519AC60E CRC64;

Query Match 33.8%; Score 662; DB 1; Length 374;  
 Best Local Similarity 37.4%; Pred. No. 5.6e-39;  
 Matches 145; Conservative 79; Mismatches 142; Indels 22; Gaps 8;

QY 1 MASLAANAEEFCNLFREMDNQGNGVFFSSLSLFAALALVFLGAGDDSLSDIKLHV 60  
 DB 1 MDDLCRANGTFAISLKLGEEDNSRNVFFSPMSISSALAMVFGKAGSTAAQWSQALCL 60  
 QY 61 NTASGYGNSNSQGLQKRVFSDINASHKDYDLSTVGLFAEKVGFHKVIECAEK 120  
 DB 61 Y-----KGDHHRFQSL-----SEVNRGTQYLLRTANRKGEXTCDPLPDFEYCK 110  
 QY 121 LYDAKVRVDFTNHLDETRNKNWENETHGKIKNVGRGGISSANWLVNAVYFKG 180  
 DB 111 FYQAELELFAEDTECRKEHINDWAEKTEKISEVLDAGTVDPLTKLVNAVYFKG 170  
 QY 181 WSAFTKSETINCHPKSCGKAVAMHQRKFNLSVIEDPQNKILELY-NGGINVY 239  
 DB 171 WNEQFDRKYTRGMLFATNE-EKKIVQMPFEAKFNGYADEVITQVLELPYVESELNVI 229  
 QY 240 LLPEN--DLSEIENKLTQNLMEWTNPRMTSKYVEFFPQFKIENYEMKQYRALGLK 297  
 DB 230 LLPDNDTDLAVVEKALTYEKKAWTNEKTKSKVQVFLPKLESDYDLEPFLRLGMI 289  
 QY 298 DIPDESADLSGTASGRLYISRMHKSIVTEVTEGTEATAATG---SNIVEKQLPOSTL 354  
 DB 290 DAFDEAKADFGSMSTKXNPLSKVAHKCFVEVNEEGTEAANAATAVNRSCRMPEPR--- 346  
 QY 355 PRADHPFLFVIR--KDDILFSGKVSCP 380  
 DB 347 FQADHPFLFIRHKNTCLIFGGRFSP 374

## RESULT 15

SE10 HUMAN  
 ID SE10 HUMAN STANDARD; PRT; 397 AA.  
 AC P48595;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Bonapin (Protease inhibitor 10) (Serpin B10).  
 GN SERPINB10 OR P110.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;

RX MEDLINE=96070759; PubMed=7592909;  
 RA Riewald M., Schlegel R.R.;  
 RT "Molecular cloning of bonapin (protease inhibitor 10), a novel human  
 RT serpin that is expressed specifically in the bone marrow.";  
 RL J. Biol. Chem. 270:26754-26757(1995).  
 CC -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF PROTEASE  
 CC ACTIVITIES DURING HEMATOPOIESIS.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE BONE MARROW.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: U35459; AAC50282.1; --  
 CC PIR: I39184; I39184.  
 CC HSP: P05619; IHLE.  
 CC Genew: HGNC:8942; SERPINB10.  
 CC MIM: 602058; --  
 CC GO: GO:0005515; F:protein binding activity; TAS.  
 CC GO: GO:0004868; F:serpin; TAS.  
 CC InterPro: IPR000215; Serpin.  
 CC Pfam: PF00079; serpin; 1.  
 CC SMART: SM00093; SERPIN; 1.  
 CC PROSITE: PS00284; SERPIN; 1.  
 CC Serpin; Serine protease inhibitor.  
 CC ACT\_SITE 362 363 REACTIVE BOND (BY SIMILARITY).  
 CC SEQUENCE 397 AA; 45402 MW; 8CE01246867154DF CRC64;

Query Match 33.6%; Score 658.5; DB 1; Length 397;  
 Best Local Similarity 34.9%; Pred. No. 1.1e-38;  
 Matches 139; Conservative 89; Mismatches 151; Indels 19; Gaps 5;

QY 1 MASLAANAEEFCNLFREMDNQGNGVFFSSLSLFAALALVFLGAGDDSLSDIKLHV 60  
 DB 1 MDSLATSINQFALESKSLAESAGQKNIFPSSWSISTSLTIVLGAKGTAAQMAQVLPQ 60  
 QY 61 NTASGYG-----NSNSQGLQKRVFSDINASHKDYDLSTVGLFAEKV 107  
 DB 61 NRDQVKCDPESKKKKEFNLSSEB-IHSDPQTLISEILKPNDDVLLKTNALVYGEKT 119  
 QY 108 YGFHKDYIECAEKLYDAKVRVDFTNHLDETRNKNWENETHGKIKNVIEGEGISSA 167  
 DB 120 YAFHNKYLEDMKTVFGAEPQPVNFVEASDQIRKIDINSWVERQTEGKIQNLPPDSVDSTT 179  
 QY 168 VMVLNAVYFKGWQSAFTKSETINCHPKSCGKAVAMHQRKFNLSVIEDPFSMKIL 227  
 DB 180 RMILYNAVYFKGWQSAFTKSETINCHPKSCGKAVAMHQRKFNLSVIEDPFSMKIL 239  
 QY 228 ELRYNG-GINNVYLLPE--NDLSEIENKLTQNLMEWTNPRMTSKYVEFFPQFKIKN 284  
 DB 240 QLYYKSRDLSSLILLPEDINGLEQLEKAITYKLNWTSADMWELYEYQLHPFKKLEDS 299  
 QY 285 YEMKQYRALGLKDIPEKADLSGIASGRLYISRMHKSIVTEVTEGTEATAATGSI 344  
 DB 300 YDLKSTLSSMGMSDAFSQSKADFGSGSARNFLSLNVFHKAFVFNQGTGAAGSGSEI 359  
 QY 345 VEKQPLPQSTLFRADHPFLFVIR--KDDILFSGKVSCP 380  
 DB 360 DIRIRVPSIEFNANHPFLFIRHKNTCLIFGGRFSP 397

Search completed: December 12, 2003, 16:25:58

Job time : 19 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 16:23:17 ; Search time 41 Seconds  
(without alignments)  
2391.707 Million cell updates/sec

Title: US-09-508-997A-2  
Perfect score: 1959  
Sequence: 1 MASAAANAEPFNLFREMD.....FLFVIRKDDIILFSGKVSQCP 380

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rviro:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query %			DB ID	Description
	Score	Match	Length		
1	1505	76.8	380	11 Q9D695	Q9D695 mus musculus
2	1488	76.0	380	11 Q920J5	Q920J5 rattus norv
3	800	40.8	388	11 Q91212	Q91212 mus musculus
4	798	40.7	388	11 Q9C0V3	Q9C0V3 mus musculus
5	763.5	39.0	423	11 Q9D7P9	Q9D7P9 mus musculus
6	734	37.5	390	4 Q81X13	Q81X13 homo sapien
7	710.5	36.3	379	11 Q8BK60	Q8BK60 mus musculus
8	708.5	36.2	379	11 Q9D154	Q9D154 mus musculus
9	707	36.1	382	11 Q8VHP7	Q8VHP7 mus musculus
10	700.5	35.8	379	11 Q9D7S8	Q9D7S8 mus musculus
11	696.5	35.6	367	11 Q8EHL1	Q8EHL1 mus musculus
12	694.5	35.5	369	4 Q9BYF7	Q9BYF7 homo sapien
13	687.5	35.1	387	11 Q8BG86	Q8BG86 mus musculus
14	686.5	35.0	387	11 Q9D1Q5	Q9D1Q5 mus musculus
15	684.5	34.9	377	11 Q08804	Q08804 mus musculus
16	684.5	34.9	385	11 Q8VHQ0	Q8VHQ0 mus musculus

Q92Z92 mus musculus  
Q73860 meleagris g  
Q8C0C0 mus musculus  
Q9D1E7 mus musculus  
Q9D1E7 mus musculus  
Q9D1E7 mus sapien  
Q08797 mus musculus  
Q8K3Y1 mus musculus  
Q9DAV6 mus musculus  
Q08800 mus musculus  
Q8BMT0 mus musculus  
Q8K1K6 mus musculus  
Q08806 mus musculus  
Q9D0S8 mus musculus  
Q8AYE3 brachydanio  
Q8VHQ1 mus musculus  
Q8K3K4 rattus norv  
Q8VUS2 struthio ca  
Q9DAZ7 mus musculus  
Q8VUS0 xenopus lae  
Q8VUS1 cheilydra se  
Q9PTA8 salmo salar  
Q8TNN7 methanosarc  
Q73790 gallus gall  
Q91422 gallus gall  
Q91422 gallus sapien  
Q9D6A7 mus musculus  
Q9WT11 cavia porce  
Q8K15 methanosarc  
Q9W648 fugu rubrip

## ALIGNMENTS

RESULT 1

Q9D695

ID Q9D695 PRELIMINARY; PRT; 380 AA.

AC Q9D695;

DT 01-JUN-2001 (TREMELrel. 17, Created)

DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)

DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)

DE 4631416M05Rik protein (Megin).

GN SERPINB7 OR 4631416M05RIK.

OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]\_taxid=10090;

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/60; TISSUE=Skin;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
RA Saito T., Okazaki Y., Gojoberi T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Stauber P., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein X.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustinich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordore P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-P.,  
RA Suzuki H., Tcyo-oka K., Wang K.H., Weitz C., Whittaker C., Whiting L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,  
RA Hayashizaki Y.,  
RL "Functional annotation of a full-length mouse cDNA collection."  
RN Nature 409:685-690(2001).

RP SEQUENCE FROM N.A.

```

PX MEDLINE=21368006; PubMed=11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Tagi M., Nagano N., Inagi R., Kurokawa K.; regulation in
RT "Cloning of rodent megins revealed its up-regulation in
RT mesangio proliferative nephritis."
RL Kidney Int. 60:641-652(2001).
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AK014524; BAB29410.1; -
DR ENBL: AF105328; AAL16768.1; -
DR HSSP; P05121; IA7C.
DR MGD; MGI:2151053; Serpinb7.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; Serpin; 1.
DR SMART; SM00093; Serpin; 1.
DR PROSITE; PS00284; Serpin; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
DR PROSITE; PS00284; Serpin; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCFB9CF4 CRC64;

Query Match 76.8%; Score 1505; DB 11; Length 380;
Best Local Similarity 73.9%; Pred. No. 2.2e-99;
Matches 281; Conservative 53; Mismatches 46; Indels 0; Gaps 0;

QY 1 MASIAAANAFCNLFREMDNQGNGNVFSSLSLPAALAVRLGAGDDSLSQIDKLIHV 60
DB 1 MASIAAANAFCNLFREMDNQGNGNVFSSLSLPAALAVRLGAGDDSLSQIDKLIHV 60
QY 61 NTASGYGNSNSQSGQLQKVFSDINASHKDYDLSIVNGLPAKVGFGHKDYIECAEK 120
DB 61 NPSRQGNSSNQSGQLQKVFSDINASHKDYDLSIVNGLPAKVGFGHKDYIECAEK 120
QY 121 LYDAKVERVDTNHELDTRNINRWENETHGKIKRVNIGEGGSSSAVMVLNAYYFKG 180
DB 121 LYDAKVERVDTNHELDTRNINRWENETHGKIKRVNIGEGGSSSAVMVLNAYYFKG 180
QY 181 WSAFTKSETINCHFKSPCKSGKAVMMHQRKFNLSVIEDPSPMKILELYNGGNNMVL 240
DB 181 WSAFTKSETINCHFKSPCKSGKAVMMHQRKFNLSVIEDPSPMKILELYNGGNNMVL 240
QY 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPFPFKIEKNYEMKYLALGDKDIF 300
DB 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPFPFKIEKNYEMKYLALGDKDIF 300
QY 301 DESKADLSGASGRLYIRMMHKSIVIEVTEGTATATGNSNIVEKQLPOSTLPRADHP 360
DB 301 DESKADLSGASGRLYIRMMHKSIVIEVTEGTATATGNSNIVEKQLPOSTLPRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2
Q920J5 PRELIMINARY; PRT; 380 AA.
AC Q920J5
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RIKEN cDNA 2310046M08 gene.
GN 2310046M08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; BC010313; AAL10313.1; -
DR HSSP; P01009; IQLP.
DR MGD; MGI:1914207; 2310046M08RIK.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; Serpin; 1.
DR SMART; SM00093; Serpin; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 388 AA; 43544 MW; AF4BB62FB084E6FA CRC64;

Query Match 40.8%; Score 800; DB 11; Length 388;
Best Local Similarity 41.8%; Pred. No. 4.5e-49;
Matches 162; Conservative 81; Mismatches 137; Indels 8; Gaps 5;

QY 1 MASIAAANAFCNLFREMDNQGNGNVFSSLSLPAALAVRLGAGDDSLSQIDKLIHV 60

```

```

DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; Serpin; 1.
DR SMART; SM00093; Serpin; 1.
DR PROSITE; PS00284; Serpin; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 380 AA; 42821 MW; DB076CA8EB2C2FBC CRC64;

Query Match 76.0%; Score 1488; DB 11; Length 380;
Best Local Similarity 73.7%; Pred. No. 3.6e-98;
Matches 280; Conservative 57; Mismatches 43; Indels 0; Gaps 0;

QY 1 MASIAAANAFCNLFREMDNQGNGNVFSSLSLPAALAVRLGAGDDSLSQIDKLIHV 60
DB 1 MASIAAANAFCNLFREMDNQGNGNVFSSLSLPAALAVRLGAGDDSLSQIDKLIHV 60
QY 61 NTASGYGNSNSQSGQLQKVFSDINASHKDYDLSIVNGLPAKVGFGHKDYIECAEK 120
DB 61 ISPSQGNSSNQSGQLQKVFSDINASHKDYDLSIVNGLPAKVGFGHKDYIECAEK 120
QY 121 LYDAKVERVDTNHELDTRNINRWENETHGKIKRVNIGEGGSSSAVMVLNAYYFKG 180
DB 121 LYDAKVERVDTNHELDTRNINRWENETHGKIKRVNIGEGGSSSAVMVLNAYYFKG 180
QY 181 WSAFTKSETINCHFKSPCKSGKAVMMHQRKFNLSVIEDPSPMKILELYNGGNNMVL 240
DB 181 WSAFTKSETINCHFKSPCKSGKAVMMHQRKFNLSVIEDPSPMKILELYNGGNNMVL 240
QY 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPFPFKIEKNYEMKYLALGDKDIF 300
DB 241 LPENDLSIENKLTFFQNLMEWTPRMTSKYVEVFPFPFKIEKNYEMKYLALGDKDIF 300
QY 301 DESKADLSGASGRLYIRMMHKSIVIEVTEGTATATGNSNIVEKQLPOSTLPRADHP 360
DB 301 VESRADLSGASGRLYIRMMHKSIVIEVTEGTATATGNSNIVEKQLPOSTLPRADHP 360
QY 361 FLFVIRKDDIILFSGKVSCP 380
DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 3
Q91Z12 PRELIMINARY; PRT; 388 AA.
AC Q91Z12
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RIKEN cDNA 2310046M08 gene.
GN 2310046M08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; BC010313; AAL10313.1; -
DR HSSP; P01009; IQLP.
DR MGD; MGI:1914207; 2310046M08RIK.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; Serpin; 1.
DR SMART; SM00093; Serpin; 1.
DR Protease inhibitor; Serine protease inhibitor; Serpin.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 388 AA; 43544 MW; AF4BB62FB084E6FA CRC64;

Query Match 40.8%; Score 800; DB 11; Length 388;
Best Local Similarity 41.8%; Pred. No. 4.5e-49;
Matches 162; Conservative 81; Mismatches 137; Indels 8; Gaps 5;

QY 1 MASIAAANAFCNLFREMDNQGNGNVFSSLSLPAALAVRLGAGDDSLSQIDKLIHV 60

```



```
RP SEQUENCE FROM N.A.
PC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
CC -I- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AK009018; BAB26028.1; -.
DR EMBL; AK040697; BAC30672.1; -.
DR HSSP; P05619; IHLE.
DR MGD; MGI:1919119; 2300003F07Rik.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin.1.
DR SMART; SM00093; SERPIN.1.
DR PROSITE; PS00284; SERPIN.1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 423 AA; 47834 MW; 5A22BE2FE51B6120 CRC64;

Query Match 39.0%; Score 763.5; DB 11; Length 423;
Best Local Similarity 37.8%; Pred. No. 2e-46;
Matches 160; Conservative 77; Mismatches 143; Indels 43; Gaps 4;

QY 1 MASLAANAEEFCFNLPREMDNQNGNVFFSSLSFAALALVRLGAGQDLSQIDKLHV 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MDSLTAANKFCDFPRE-SKDDAHKHFVCPLSLSAAGMVLGARGDSAHQIDEALHF 60
QY 61 NTAS-----GYGNSNSQSGLOQLARVPSD 86
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 NELSDEHKEPNDFPQSPSKASDSSLEGOKQTASQDQGGESTNDHQLGCHFGKLSLR 120
QY 87 INASHKDYLSIVNGLFAEKYGFHKDYICAEKLYDAKVRVDFTNHLEDTRRNINRW 146
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 IPRDSXYTLWMANLYGQEPPICSHYSDDVTEPHHTTVESVDFQDSEKSRQENFW 180
QY 147 ENETHGKIKNVIGRGSSAVMVLVNAVYFKGWSAFTKSETINCHFKSPCKSGKAVA 206
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 ESQSQKIKELFGKEALDNSTVLVNAVYFKAWEREPNSENTVDASFLCNEKTKVK 240
QY 207 MHQBEKKNLNVIEDPSMKILELRY-NGGINVYLLPE-----NDLSEIKNLIFQNL 259
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 MNQKGRFRIGFIDELQAOILEMKYAMGKLSMLVLLPSCSDNNVNSQLEKKINHEKL 300
QY 260 EWTNPRMTSKYVEFPQFKIEKXVEMKQYLKALGLKIDFDESKADLSGIASGRLYIS 319
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 ANSSSENLSKFPVLSFPQNLDSYDLKSLIQDMGIKDVFDETKADLTGSKSPNLYLS 360
QY 320 RMHKSIEVTEGTETATATGSNIVEKQLPQSTLFRADHPFLFVIRKD--DILLPSGV 377
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 KIVHKTFVEVDENGTQAAASGVVAEXALPSEWFEANHPFLFFIRINPTQSLPFCGRV 420
QY 378 SCP 380
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 YCP 423

RESULT 6
Q8IXI3 PRELIMINARY; PRT; 390 AA.
ID Q8IXI3
AC Q8IXI3
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Squamous cell carcinoma antigen 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore P.L., Ong S., Harrison T.J.;

Query Match 36.3%; Score 710.5; DB 11; Length 379;
Best Local Similarity 39.1%; Pred. No. 1.1e-42;
Matches 152; Conservative 74; Mismatches 144; Indels 19; Gaps 6;

QY 1 MASLAANAEEFCFNLPREMDNQNGNVFFSSLSFAALALVRLGAGQDLSQIDKLHV 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MEQJSSANTLFALELDFQTLNESPFTGNIFPSPFSSSALAMVLGKSGTAQLSKTFH- 59
QY 61 NTASGYGNSNSQSGLOQLARVPSDINASHKDYLSIVNGLFAEKYGFHKDYICAEK 120
"Binding of HBV to cells is mediated by SCCa1 but does not require the
RT reactive site loop.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ515706; CAD56658.1; -.
SQ SEQUENCE 390 AA; 44504 MW; E9D56D2D786C9E24 CRC64;

Query Match 37.5%; Score 734; DB 4; Length 390;
Best Local Similarity 40.9%; Pred. No. 2.3e-44;
Matches 164; Conservative 80; Mismatches 125; Indels 32; Gaps 11;

QY 1 MASLAANAEEFCFNL---PREMDNQNGNVFFSSLSFAALALVRLGAGQDLSQIDKL 57
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MMSLSEANTKEMFDLSQOFRKSEN---NIFVPSITSALGMVLGAKNTAQOIKKV 56
QY 58 LHV-----NTASGYGNSNSQSG-LSQQLARVPSDINASHKDYLSIVNGLFAEKYGFH 111
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 LHPDQWENTTGAATYHYDRSGNVHQFQKLTTEFNKSTDAYELKIANKLFGEKTYLFL 116
QY 112 KQYISCAKLYDAKVRVDFTNHLEDTRRNINKEWVENETHGKIKNVIGEGGSSAVMVL 171
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 QEYLDALIKKFYQTSVSDVFANAPESGRKKINWSVESQNEKIKNLIPENIGSNTLVL 176
QY 172 VNAVYFKGWSAFTKSETINCHFKSPCKSGKAVAMHQERKFNLSVIEDPSMKILELRY 231
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
177 VNAVYFKGWSAFTKSETINCHFKSPCKSGKAVAMHQERKFNLSVIEDPSMKILELRY 236
QY 232 NG-GINMYVLLPE--NDLSEIKNLIFQNLMEWTNPRMTSKYVEFPQFKIEKXVEMK 288
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 KGKDSMIVLLPNEIDGLQKLEKLTAKLMEWTSIQNNRETRVLDLHPFKVEBSYDK 296
QY 289 QYLKALGLKIDFDESKADLSGIASGRLYISRMHKSIEVTEGTETATAT-----GSN 343
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 DTLRWNGVDLIFN-GDADLSGMTSGRGLVLSGLVHARFVETESGAEAAATAVAFSS 355
QY 344 IVEKQLPOST--LFRADHPFLFVIR--KDDILFSGKVSPP 380
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 -----PSTWEEPHCHNPFLFFIRQNTNLSILFYGRFPSP 390

RESULT 7
Q8BK60 PRELIMINARY; PRT; 379 AA.
ID Q8BK60
AC Q8BK60;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Serine.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK076129; BAC36206.1; -.
SQ SEQUENCE 379 AA; 42618 MW; D6B1DB7B6EBC845C5 CRC64;

Query Match 36.3%; Score 710.5; DB 11; Length 379;
Best Local Similarity 39.1%; Pred. No. 1.1e-42;
Matches 152; Conservative 74; Mismatches 144; Indels 19; Gaps 6;

QY 1 MASLAANAEEFCFNLPREMDNQNGNVFFSSLSFAALALVRLGAGQDLSQIDKLHV 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MEQJSSANTLFALELDFQTLNESPFTGNIFPSPFSSSALAMVLGKSGTAQLSKTFH- 59
QY 61 NTASGYGNSNSQSGLOQLARVPSDINASHKDYLSIVNGLFAEKYGFHKDYICAEK 120
```

a greatly expanded Ov-serpin gene repertoire in the mouse.";

Genomics 0:0-0(2002).

[4]

SEQUENCE FROM N.A.

STRAIN=129/SvEvTac; TISSUE=Spleen;

RC Benarafa C., Cooley J., Zeng W., Bird P.I., Remold-O'Donnell E.;

RA Kalsnerman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,

RA McKesone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,

RA Ragoussis J., Bird P.I.;

RT "Comparison of human chromosome 6p25 with mouse chromosome 13 reveals

RT a greatly expanded ov-serpin gene repertoire in the mouse.";

RT Genomics 79:349-362(2002).

CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

RL

DR EMBL; AK003930; BAB33079.1; -

DR EMBL; BC011140; AAH11140.1; -

DR EMBL; AF426024; AAL57486.1; -

DR EMBL; AF521697; AAU59533.1; -

DR HSSP; P05619; 1HLE.

DR

DR MGD; MGI:1913472; Serpinb1a.

DR InterPro; IPR00215; Serpin.

DR Pfam; PF00079; serpin; 1

DR PROSITE; PS00284; SERPIN; 1.

DR

KW Process; Protease inhibitor; Serine protease inhibitor; Serpin.

KW

SQ SEQUENCE 379 AA; 42575 MW; F492F1EBB8E80B1 CRC64;

Query Match 36.3%; Score 708.5; DB 11; Length 379;

Best Local Similarity 37.3%; Pred. No. 1.5e-42;

Matches 145; Conservative 81; Mismatches 144; Indels 19; Gaps 4

Qy 1 MASLAANAACFCNLFREMDNQGNGNVFFSSLSFAALALVRLGAGDQSLSDIKLLHV 60

Db 1 MEQLSSANTFLALELFTLNESSPTGNIFFSPSISALLAMVILGAKGTAQAQLSKTFHF 60

Qy 61 NTASGNGSNSQSGLSQIKRVPDSINASHKDYDLSVINGLFAEKYGFHDKDYTCABK 120

Db 61 -----DSVEDTHSRFQSLNAEVSKGASHTLKANELYGEKTYAPLPEYLAATQK 110

Qy 121 LYDAKVERVDTNHLEDTNRNINKVNEETHGKIKNVEGGGTSASAVMLVNAVYFKGK 180

Db 111 MYGADLA PVDLHASEDARKEINQWKGQTEGKIPELLSVGVDSTKLVNAVYFKGM 170

Qy 181 WQAFKTSGETINCHFKSPCKSGAVAMHQRKFNLSVIEDPSMKLLELYNGG-INMYV 239

Db 171 WEKFWETDITDAPFLRLSKDQTKTVQWYQKKKFPFGVYISDLKCKVLEMPYQGGELSMVI 230

Qy 240 LLPEN-----DSEIENKLTQNLMEWTNPRAWTSKKVEYFPQPKIKENYEMKQYLRA 293

Db 231 LLPKQIDESTGLKKLEKQITLEKLLRWTKRENLEBFIDVHVKLFRFKISEYTLNSLGR 290

Qy 294 LGLKQIDPESKADLSGIAGSGGLYTSRMHKSIVETEGTEGTEATAATGSNIVEKQLPOST 353

Db 291 LGVQDLFSKADLSGNSGSRDLFTSKIVKSFVNEEGTEAATAATGATFCMLPEE 350

Qy 354 LFRADHPFLFVIRKDDI--ILFSGKVGSCP 380

Db 351 EFTVDHPPIFFIRHNTSNVLFLGRVCSF 379

RESULT 9

Q8VHP7

ID Q8VHP7 PRELIMINARY; PRT: 382 AA.

AC Q8VHP7;

DT C1-MAR-2002 (trEMBLrel. 20, Created)

DT C1-MAR-2002 (trEMBLrel. 20, Last sequence update)

DT C1-MAR-2003 (trEMBLrel. 23, Last annotation update)

DE Serine protease inhibitor E1B (Serine (or cysteine) proteinase



Db 111 MYGAADLPVDFHASEDARKEINQWVKQTEGKIPELLSVGVDSMTKLVLVNAIYFKGM 170  
 QY 181 WQSAATKSETINCHFKSPKCSKAVAYMHQERKFNLSVIEDPSMKILEURYNGG-INYYV 239  
 Db 171 WEEKWNTEDTDAFFLSKQDTRVYKMYQKXPGYISDLCKKVLMEPYQGGELSMVI 230  
 QY 240 LLPEN-----DASEENKLTQNLMEWTPRMTSKYVEVFPQPKIEKMYEQYLRA 293  
 Db 231 LPPKIDIESNGLKLEKQILEKLEWTRENFLEIDVHVKLPFKIESYLSNIGR 290  
 QY 294 IGLKIDFDSKADLSGIASGRLYISRMHKSIVETEGTEATAATGNSIVEXQLPOST 353  
 Db 291 LGVQDLFESSKADLSGWSGRDLFTSKIVHKSFEVNEEGTEAAATGGIATFCMLPDEE 350  
 QY 354 LFRADHPFLFVKIDLI--ILFSKVSQCP 380  
 Db 351 EFTVDHPFFIFIRHNPTSNVLFLGRVCSQ 379

RESULT 11  
 Q9BYF7 PRELIMINARY; PRT; 387 AA.  
 ID Q9BYF7  
 AC Q9BYF7  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DE Squamous cell carcinoma antigen 2 related protein 1.  
 GN SCQA2-RS1  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Car-S, and Car-R; TISSUE=TPA-treated skin;  
 RA Gariboldi M., Peissel B., Fabbri A., Saran A., Zaffaroni D.,  
 RA Spinola M., Falvella P.S., Pazzaglia S., Tanuma J.-I., Maurichi A.,  
 RA Bartoli C., Silverman G., Coveilli V., Pilotti S., Hayashiaki Y.,  
 RA Okazaki Y., Dragani T.A.;  
 RT "The serpin Scqa2 gene plays a functional role in genetic  
 RT susceptibility to skin tumorigenesis in mice and humans."  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY144685; AAN62872.1; -  
 DR EMBL; AY144686; AAN62873.1; -  
 SQ SEQUENCE 367 AA; 44539 MW; 0AD6398FA0251135 CRC64;

Query Match 35.6%; Score 696.5; DB 11; Length 387;  
 Best Local Similarity 37.3%; Pred. No. 1.1e-41;  
 Matches 145; Conservative 91; Mismatches 142; Indels 11; Gaps 6;

QY 1 MASLAANAARFCNLFREMDNQGNGVVFSSLSLFAALALVRLGAQDLSQIDKLH- 59  
 Db 1 MIRHAADKFAVEMYQL--RSDKNIFYSPISTMTALQLAGAKNTGIEKVIQF 58  
 QY 60 VNTASGVNSS--NSQSGLOSQKRVFSDINASHKDYDLSVNGLFAEKVYGVGHKDY 116  
 Db 59 IETTKTKTKSEHCDDEENVHEQFKLITQLKNSDNDYDLKAANSYGAQKFPFLQFLE 118  
 QY 117 CAELKYDAKVERVDFTNHLBTRNINKNWENTHKGINKVIGEGGISSAVMLVNAVY 176  
 Db 119 DIKXYQAKVSELDPEHATESERKINSVSKTNGIKDLPPSGSLSSITLVNAVY 178  
 QY 177 FKGWQSAFTKSETINCHFKSPKCSKAVAYMHQERKFNLSVIEDPSMKILEURYNG- 235  
 Db 179 PKGWNKXEDNHTREKFWLNKNTSKPYQMMKQKRNKFNFSFLGDVHAQIVSEIYKQKOL 238  
 QY 236 NMYVLPE--NDLSIENKLTQNLMEWTPRMTSKYVEVFPQPKIEKMYEQYLRA 293  
 Db 239 SMFVLLPMEIDGKLEQLQTTDLKLENTKAENMELTLYLSLPRFKVEEKYDQVPLSH 298  
 QY 294 LGLKIDFDSKADLSGIASGRLYISRMHKSIVETEGTEATAATGNSIVEXQLPOST 353  
 Db 299 MGWVDAFPDQKADFGSGMSSIPGLWSKVLHKSFEVNEEGTEAAATGVEVSQAIAE 358

QY 354 LFRADHPFLFVI--RKDDIILFSKVSQCP 380  
 Db 359 DFCDHPFLFIIHRKTNLSLFFGRISSP 387

RESULT 12  
 Q9BYF7 PRELIMINARY; PRT; 369 AA.  
 ID Q9BYF7  
 AC Q9BYF7  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE SCCA2b.  
 GN SCCA2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Suminami Y., Kishi F., Murakami A., Sakaguchi Y., Kato H.;  
 RT "Novel Forms of SCC Antigen Transcripts Produced by Alternative  
 RT Splicing."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR EMBL; AB046400; BAB40773.1; -  
 DR HSSP; P01008; 1ATH.  
 DR InterPro; IPR000215; Serpin.  
 DR Pfam; PF00079; Serpin; 1.  
 DR SMART; SM00093; Serpin; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 KW Protease inhibitor; Serine protease inhibitor; Serpin.  
 SQ SEQUENCE 369 AA; 42287 MW; CED7A2426378DE3E CRC64;

Query Match 35.5%; Score 694.5; DB 4; Length 369;  
 Best Local Similarity 39.0%; Pred. No. 1.4e-41;  
 Matches 153; Conservative 82; Mismatches 122; Indels 35; Gaps 10;

QY 1 MASLAANAARFCNLFREMDNQGNGVVFSSLSLFAALALVRLGAQDLSQIDKLH- 60  
 Db 1 MNSLSEANTKPMDFLPQQRKSKEN-NIFYSPISITTSALGMVLGAKDNTAQQISKVLHF 59  
 QY 61 -----NTASGVNSSQSG-LQSLKRVFSDINASHKDYDLSVNGLFAEKVYGVGHKDY 114  
 Db 60 DQWTEENTKAAATYHVDRSGNVHGFQKLLTETNKSTDAYELKIANKLFGKTYQFLQY 119  
 QY 115 IECAELKYDAKVERVDFTNHLBTRNINKNWENTHKGINKVIGEGGISSAVMLVNA 174  
 Db 120 LDALKKFFQTSVSESTDFAAPEERKKINSWVESQTNKIKNLPDGTIGNDTLVLVNA 179  
 QY 175 VFYKGWQSAFTKSETINCHFKSPKCSKAVAYMHQERKFNLSVIEDPSMKILEURYNG- 233  
 Db 180 IYFKQWENPKENT-----KEEFWEN--KDVQAKVLEIPIYKKG 218  
 QY 234 GINMYVLPE--NDLSIENKLTQNLMEWTPRMTSKYVEVFPQPKIEKMYEQYL 291  
 Db 219 DLSMIVLLPNEIDGLQKLEELTAELMEWTSIQNRETCDLHLPRFWESYDLDKTL 278  
 QY 292 RALGKIDFDSKADLSGIASGRLYISRMHKSIVETEGTEATAATGNSIVEXQLPQ 351  
 Db 279 RTMGWVAIFN-GPADLSGMTWSHGLSVSKVLHKAPEVTEEGVEAAATAVAVVVELSSPS 337  
 QY 352 SP-LFRADHPFLFVI--KDDIILFSKVSQCP 380  
 Db 338 TNEEFCCNHEFFLFFIRQNKTNLSLFFGRFSSP 369

RESULT 13  
 Q9BG86 PRELIMINARY; PRT; 387 AA.  
 ID Q9BG86  
 AC Q9BG86  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakura I.,  
RA Saito T., Okazaki Y., Gofobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischnann W., Gaasterland T., Giasi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikiado I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,  
RA Blake J., Focellì D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustinchin S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC !- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
DR EMBL; AK003220; BAB2850.1; -.  
DR HSSP; P05120; 1BY7.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00079; serpin.1.  
DR SMART; SM00093; SERPIN.1.  
DR PROSITE; PS00284; SERPIN.1.  
KW Protease inhibitor; Serine protease inhibitor; Serpin.  
SQ SEQUENCE 387 AA; 44558 MW; AE25C0520F3B184 CRC64;  
  
Query Match 35.0%; Score 686.5; DB 11; Length 387;  
Best local similarity 36.8%; Pred. No. 5.6e-41;  
Matches 143; Conservative 91; Mismatches 144; Indels 11; Gaps 6;  
  
QY 1 MASLAANAAGFCNLPREMDDNQNGVFFSLSLFLALALVELGAQDDLSQIDKLH- 59  
Db 1 MTRFFAADVFAVMYRQL--RESDKNIFFYSPISMTALMLQLGAKGNTIEIKVLQF 58  
QY 60 VNTASGYGNSS--NSQGLQSLKQKVFDDKASHKDYDLSVGLFAEKYVGFHKDYIE 116  
Db 59 IETTKTEKSEHCDDEENVEHQFKLITQLNKSNDYDYLKAANSIYAKGFFPFLQFLE 118  
QY 117 CAEKLYDKAVRVDFTNHLRETRINKVWENHKGIRNVIGEGISSSAWVLVNAVY 176  
Db 119 DKEYQAKVESIDPHEATESEKINSWVEKTINGKIDLPFGSLSSSTILVLVNAVY 178  
QY 177 FKGQKQSAFTKFTFNCIFKSPKSGKAVMMHDERKFNLSVIEDPSMKILELRYNG-GI 235  
Db 179 FGQWRKRFKVENTREKFWLNKNTSKPQMKNQENFNSFLGDVHAQIVEIPYKGLD 238  
QY 236 NYVLLPE--NDLSELENKLTFTQNLMEWTPNPRMTSKYVEVFPFQFKIKNRYEMKQYLA 293  
Db 239 SMFVLLPEIDGLKLEHQLTITDKLELWIKASNNHLTELYLSLPFRKVEEKYDLQVPLEH 298  
QY 294 LGLKQIDPESKADLIGTASGGRVLSRWVHKSYEVTEEGTGTATATCGSNIIVEKQLPQST 353  
Db 299 MGNVADPDPQADFGSSNSIGELVSKVLHKSFEVNEEGTEAAATGVEVSRSAQIAE 358  
  
QY 354 LFRADPEPLFVY--RKDDIIILFSKGVSCP 380  
Db 359 DFCDDHPLFPIHRTWNSILFFGRICSP 387  
  
RESULT 15  
ID O08804 PRELIMINARY; PRT; 377 AA.  
AC O08804;  
DT 01-JUL-1997 (TReVBurel. 04, Created)  
DT 01-DEC-2001 (TReVBurel. 19, Last sequence update)  
DT 01-MAR-2003 (TReVBurel. 23, Last annotation update)  
DE Serine protease inhibitor NK13 (Serine protease inhibitor 12).  
OS SP112.  
GN Mus musculus (Mouse).



Search completed: December 12, 2003, 16:26:52  
Job time : 43 secs

Query Match	34.9%	Score 684.5	DB 11	Length 377
Best Local Similarity	39.4%	Pred. No. 7.6e-41		
Matches 152	Conservative 76	Mismatches 143	Indels 15	Gaps 8
QY	1	MASLAAANAECECNLFREMDNQNGQNVFSSLSIFPALALVRLGAQDSSLSQIDKLHV	60	
Db	1	MDPLEANNFALNLLKTUGDSSR-NVLFPSPISVSALAMVFMGAKTTASQMAQALS	59	
QY	61	NTASGYGNSNSQSGLOSQLKRVFSDINASHKDYDLISVNGLFAEKYVGHKDYTECAEK	120	
Db	60	DKCSGKGGRDVHQ-QFOSLL-----TEINTGTQVYLTANLPLFGKTFDILASFKDCSRK	114	
QY	121	LYDAKVERVQFTWHLEDTRNIKNWENTHCKIKWIGEGISSAWMYLVNAVYFKGK	180	
Db	115	FIERMEELFOFKATEQGRQHINAWAKYDEKITELLSSGSWNSNTPLVNAVYFKGN	174	
QY	181	WOSAFKTSETINCHFKSPKCSKAVAMHQRKCNLSVIEDPSMKILELYNGG-INMYV	239	
Db	175	WEKQNKEDQEMPFNVKQVYFVQWQMKSTFKMTYVEEISTNIIILLPYGVNSLNMI	234	
QY	240	LIPEN-DLSEINKLTQNIEMWNTPRMTSKYVEYFFQPKIEKNYEMKQYLSALGK	297	
Db	235	MIPDEHIELSMVEKBITTYKFEIWRDLXMEEBEVEVFLPKFLEENYDMQVLCRLGNT	294	
QY	298	DIFDESXADLSGIFAGGRLYXISRMWKSIVETBGTCTEATAATGSNI-VEKQLPOSTLFR	356	
Db	295	DAFEEGVADPSGATSKGELFLSKVHHKGFVWNEEGTEAAATAATGFRCMVP---	351	
QY	357	ADHPFLFVIR--KDDILFSGKVSQCP	380	
Db	352	ANHPFLFFQHSRTSGIVFCGRFSSP	377	